

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds
(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-14

Perfect score: 45

Sequence: 1 VLYLNRKGI 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/abs/ABSSWEB/spool/US10774176/runat_24042006_165112_19185/app_query.fasta.1
-DB=n_Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05p
-USER=US10774176 -CGEN 1.1 3463 @runat_24042006_165112_19185 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 21.1*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*
14: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	45	100.0	246	10	ADK11641	Adk11641 Breast ca
2	45	100.0	475	13	ADU11677	Adu11677 Solid tum
3	45	100.0	901	3	AAA27060	Aaa27060 Canine 5T
4	45	100.0	927	6	ABT07721	Abt07721 Breast ca

5	45	100.0	927	8	ABX76333	Abx76333 Lung canc
6	45	100.0	927	10	ADB80503	Adb80503 Ovarian c
7	45	100.0	927	11	ADN38723	Adn38723 Cancer/an
8	45	100.0	973	8	AAD56198	Aad56198 Human LRR
9	45	100.0	1156	6	ABV99349	Abv99349 Human NOV
10	45	100.0	1260	6	ABK87175	Abk87175 cDNA enco
11	45	100.0	1260	10	ADB97513	Adb97513 Feline 5T
12	45	100.0	1260	10	ADB97452	Adb97452 DNA enco
13	45	100.0	1263	3	AAA27058	Aaa27058 Human 5T4
14	45	100.0	1263	4	AAF89736	Aaf89736 Nucleotid
15	45	100.0	1263	6	ABK87174	Abk87174 cDNA enco
16	45	100.0	1281	3	AAA27059	Aaa27059 Mouse 5T4
17	45	100.0	1331	8	AAD56199	Aad56199 Human LRR
18	45	100.0	2020	10	ADJ56299	Adj56299 Human CDN
19	45	100.0	2053	8	ACC51052	Acc51052 Human bla
20	45	100.0	2053	8	ABX76332	Abx76332 Lung canc
21	45	100.0	2053	8	AAD56197	Aad56197 Human LRR
22	45	100.0	2053	8	AAD56200	Aad56200 Human LRR
23	45	100.0	2053	11	ADN38721	Adn38721 Cancer/an
24	45	100.0	2053	12	ADL06473	Adl06473 Human tum
25	45	100.0	2053	12	ADN03961	Adn03961 Antipori
26	45	100.0	2053	13	ADR25444	Adr25444 Breast ca
27	45	100.0	2053	13	ACN38510	Acn38510 Tumour-as
28	45	100.0	2053	13	ADV35098	Adv35098 Human CDN
29	45	100.0	2338	5	AAS87175	Aas87175 DNA enco
30	45	100.0	2359	4	AAK94253	Aak94253 Human ful
31	45	100.0	2359	12	ADL30831	Adl30831 Full leng
32	45	100.0	2361	4	AAK94254	Aak94254 Human ful
33	45	100.0	2361	12	ADI26162	Adi26162 Human CDN
34	45	100.0	2361	12	ADL30833	Adl30833 Full leng
35	45	100.0	2557	12	ADI26160	Adi26160 Human CDN
36	45	100.0	2557	12	ADI26158	Adi26158 Human CDN
37	42	93.3	1149	6	ABA90338	Ab90338 Human pol
38	42	93.3	3993	13	ADU01816	Adu01816 Novel hum
39	42	93.3	4289	4	ABL11890	Ab111890 Drosophil
40	42	93.3	4289	4	AAS57114	Aas57114 DNA enco
41	42	93.3	4289	10	ADC35812	Adc35812 Drosophil
42	41	91.1	874	13	ADX27916	Adx27916 Plant ful
43	41	91.1	110000	13	ABD32791_2	Abd32791 Plant ful
44	39	86.7	99291	13	ABD33549	Abd33549 Human can
45	39	86.7	114596	14	ADZ70593	Adz70593 Human CDN
46	39	86.7	157875	6	ABK99972	Abk99972 Human CAD
47	38	84.4	1209	6	ABK75509	Abk75509 Bacillus
48	38	84.4	1859	6	ABQ33830	Abq33830 Oligonuc
49	38	84.4	1859	6	ABQ33831	Abq33831 Oligonuc
50	38	84.4	3617	6	ABQ71002	Abq71002 Listeria
51	38	84.4	8509	4	AAI67035	Aai67035 Nucleotid
52	38	84.4	8509	6	AAS18307	Aas18307 DNA seque
53	38	84.4	290547	13	ABD32598	Abd32598 Human can
54	37	82.2	620	5	AAS05592	Aas05592 Mammalian
55	37	82.2	1077	4	ABL02409	Ab102409 Drosophil
56	37	82.2	1337	3	AAC42423	Aac42423 Arabidops
57	37	82.2	2044	13	ADT15200	Adt15200 Plant CDN
58	37	82.2	2134	10	ADB62583	Adb62583 Human CDN
59	37	82.2	2922	10	ADB62817	Adb62817 Human CDN
60	37	82.2	3140	4	ABL02408	Ab102408 Drosophil
61	37	82.2	3410	5	AAH46273	Aah46273 Human EF-
62	37	82.2	3829	5	AAH42074	Aah42074 Human cal
63	37	82.2	4174	10	ACC68992	Acc68992 Human neu
64	37	82.2	4824	12	ADQ22371	Adq22371 Human sof
65	37	82.2	40681	6	ABA92787_6	Ab92787 Human sof
66	37	82.2	63155	10	ADC85996	Adc85996 Human GPC
67	37	82.2	249487	6	ABN85733	Abn85733 Mouse gen
68	37	82.2	302603	11	ADP75187	Adp75187 Human End
69	37	82.2	330973	11	ACN44846	Acn44846 Human gen
70	36	80.0	273	12	ADL03915	Adl03915 DNA enco
71	36	80.0	467	12	ADP93371	Adp93371 Cotton ex
72	36	80.0	594	11	ACN86641	Acn86641 Breast ca
73	36	80.0	1349	2	AAV75022	Aav75022 Staphyloc
74	36	80.0	1381	11	ACN89437	Acn89437 Breast ca
75	36	80.0	1385	3	AAZ36831	Aaz36831 cDNA enco
76	36	80.0	1401	10	ACF70651	Acf70651 Phototrab
77	36	80.0	1662	13	ADT19226	Adt19226 Plant CDN

78	36	80.0	1879	10	AD28220	Ad28220 Human MDD	C 151	35	77.8	3627	13	ACN38166	Acn38166 Tumour-s8	
79	36	80.0	2095	10	ADA52975	Ada52975 Human cod	C 152	35	77.8	4113	4	ABK43507	Abk43507 DNA encod	
c	80	80.0	2223	14	ADZ64049	Adz64049 Human can	C 153	35	77.8	4113	12	ADI53894	Adi53894 cDNA enco	
81	36	80.0	2315	4	ABL24224	Abi24224 Drosophil	C 154	35	77.8	4330	6	ABN95863	Abn95863 Gene #236	
82	36	80.0	2361	10	ADD43704	Ad43704 Bacillus	C 155	35	77.8	4330	9	ACD06184	AcD06184 Human cDN	
c	83	80.0	2491	12	ADQ85388	Adq85388 Human tum	C 156	35	77.8	4330	12	ADN95977	Adn95977 Human NOV	
84	36	80.0	3375	9	ADA02809	Ada02809 Mouse Fis	C 157	35	77.8	4399	4	ABL28796	Abi28796 Drosophil	
c	85	80.0	3375	10	ADB72547	Adb72547 Mouse Fis	C 158	35	77.8	5504	5	AAS70772	Aas70772 DNA encod	
c	86	80.0	3375	10	ADC85289	Adc85289 Mouse Fis	C 159	35	77.8	5504	9	ADA02811	Ada02811 Human Fis	
c	87	80.0	3375	12	ADM74404	Adm74404 Murine ca	C 160	35	77.8	5504	10	ADB72549	Adb72549 Human CA	
c	88	80.0	3728	9	ADA02808	Ada02808 Mouse Fis	C 161	35	77.8	5504	10	ADM74406	Adm74406 Human m8N	
c	89	80.0	3728	10	ADC85288	Adc85288 Mouse Fis	C 162	35	77.8	5504	12	ADM74406	Adm74406 Human car	
c	90	80.0	3728	10	ADC85288	Adc85288 Mouse Fis	C 163	35	77.8	5504	14	ADZ09665	Adz09665 Human bre	
c	91	80.0	6173	14	ADM74403	Adm74403 Murine can	C 164	35	77.8	5521	4	ABL28840	Abi28840 Drosophil	
c	92	80.0	16463	4	ABL06068	Abi06068 Drosophil	C 165	35	77.8	5588	4	AAS22479	Aas22479 Human cDN	
c	93	80.0	28564	10	AD663335	Ad663335 Human gen	C 166	35	77.8	6077	6	ABL34218	Abi34218 Human imm	
94	36	80.0	3069	14	ADM00030	Adm00030 Lactobaci	C 167	35	77.8	6240	5	AAS90138	Aas90138 DNA encod	
c	95	80.0	99001	13	ADR20457	Adr20457 Human SPR	C 168	35	77.8	7577	4	AAS57480	Aas57480 Human liv	
c	96	80.0	103665	12	ADQ97703	Adq97703 Human can	C 169	35	77.8	11045	4	AAS46272	Aas46272 DNA encod	
c	97	80.0	110000	10	ACF67367_37	Continuation (38 o	C 170	35	77.8	11670	6	ABL54326	Abi54326 Chemicall	
c	98	80.0	110000	10	ACF65388_10	Continuation (11 o	C 171	35	77.8	13435	4	AAL199253	Aal199253 Human exp	
c	99	80.0	110000	10	ACF65388_10	Continuation (11 o	C 172	35	77.8	13435	4	AAL04187	Aal04187 Human rep	
c	100	80.0	34980	5	AH411226	Aah41226 Pyrococu	C 173	35	77.8	13435	5	AAL163603	Aal163603 Human kid	
c	101	35	77.8	130	2	AAT20898	Ata20898 Human gen	C 174	35	77.8	14704	4	AAS85044	Aas85044 Human imm
c	102	35	77.8	149	13	ACN49288	Acn49288 Cotton pr	C 175	35	77.8	15381	10	ADCO1196	Adco1196 Enterohae
c	103	35	77.8	308	13	ACN49288	Acn49288 Cotton pr	C 176	35	77.8	15393	9	ACD19178	AcD19178 E. coli 0
c	104	35	77.8	364	8	ABX49513	Abx49513 Bovine BS	C 177	35	77.8	20633	2	AAX13213	Aax13213 Enterococ
c	105	35	77.8	415	6	ABN96629	Abn96629 Gene #312	C 178	35	77.8	20633	6	ABN99008	Abn99008 Enterococ
c	106	35	77.8	466	4	AAS86624	Aas86624 Human foe	C 179	35	77.8	22585	4	AAL04299	Aal04299 Human rep
c	107	35	77.8	466	4	AAL38296	Aal38296 Probe #69	C 180	35	77.8	22927	4	AAL04782	Aal04782 Human tes
c	108	35	77.8	466	4	AAK32462	Aak32462 Human bon	C 181	35	77.8	22927	4	ABL97677	Abi97677 Human rep
c	109	35	77.8	466	4	AAK06751	Aak06751 Human bra	C 182	35	77.8	24400	14	ABE96517	Abe96517 Human IL4
c	110	35	77.8	466	4	ABS32168	Abs32168 Human liv	C 183	35	77.8	33303	3	AAA81514	Aaa81514 N. mening
c	111	35	77.8	466	6	ABS07246	Abs07246 Human gen	C 184	35	77.8	33316	12	ADQ97549	Adq97549 Human can
c	112	35	77.8	473	6	ABL56162	Abi56162 Maize acy	C 185	35	77.8	36800	10	ADB74390	Adb74390 Mycobacte
c	113	35	77.8	546	13	ADT04639	Adt04639 Microorga	C 186	35	77.8	48680	11	ACN45210	Acn45210 Human gen
c	114	35	77.8	588	12	ACH75289	Ach75289 Human gen	C 187	35	77.8	77941	11	ACN45210	Acn45210 Human gen
c	115	35	77.8	700	13	ADX50523	Adx50523 Plant ful	C 188	35	77.8	89378	12	ADN46845_20	Adn46845_20
c	116	35	77.8	700	4	AH92311	Aah92311 Human inf	C 189	35	77.8	89378	12	ADN46123_20	Adn46123_20
c	117	35	77.8	700	4	AH92311	Aah92311 Human inf	C 190	35	77.8	89378	12	ADN46464_20	Adn46464_20
c	118	35	77.8	981	8	ACN27720	Acn27720 Prokaryot	C 191	35	77.8	95001	12	ADH56439	Adh56439 Human hyp
c	119	35	77.8	991	4	ABK43824	Abk43824 DNA encod	C 192	35	77.8	110000	2	AAV21209_01	AAv21209_01
c	120	35	77.8	991	5	ADM19668	Adm19668 Novel hum	C 193	35	77.8	110000	3	AAA81489_0	AAA81489_0
c	121	35	77.8	991	12	ADI54211	Adi54211 cDNA enco	C 194	35	77.8	110000	6	ABQ69245_15	ABq69245_15
c	122	35	77.8	1197	6	ABQ90384	Abq90384 M. capsul	C 195	35	77.8	110000	6	ABQ67195_1	ABq67195_1
c	123	35	77.8	1245	10	ADG28800	Adg28800 Bacterial	C 196	35	77.8	110000	12	ADN47591_00	Adn47591 Thermococ
c	124	35	77.8	1245	12	ADP96565	Adp96565 Escherich	C 197	35	77.8	110000	12	ADN47209_00	Adn47209 Thermococ
c	125	35	77.8	1264	12	ADP96566	Adp96566 Escherich	C 198	35	77.8	110000	12	ADN47960_00	Adn47960 Thermococ
c	126	35	77.8	1500	13	ADV96728	Adv96728 Gene of t	C 199	35	77.8	110000	13	ADS99457_2	AdS99457_2
c	127	35	77.8	1509	8	ACA41314	Aca41314 Prokaryot	C 200	35	77.8	110000	14	ABE39172_1	Abe39172_1
c	128	35	77.8	1509	10	ABZ41008	Abz41008 N. gonorr	C 201	35	77.8	110000	14	ABE39175_11	Abe39175_11
c	129	35	77.8	1512	8	ACA41913	Aca41913 Prokaryot	C 202	35	77.8	110000	14	ABE42401_10	Abe42401_10
c	130	35	77.8	1628	13	ADX30480	Adx30480 Plant ful	C 203	35	77.8	110000	14	ABE42401_11	Abe42401_11
c	131	35	77.8	1850	10	ADF38058	Adf38058 Synchroni	C 204	35	77.8	110000	14	ABE42736_4	Abe42736_4
c	132	35	77.8	2109	13	ADK47566	Adk47566 Bacterial	C 205	35	77.8	128963	12	ADQ97110	AdQ97110
c	133	35	77.8	2112	3	AAZ93713	Aaz93713 F-box pro	C 206	35	77.8	137908	11	ADP65634	AdP65634
c	134	35	77.8	2183	4	ABL28797	Abi28797 Drosophil	C 207	35	77.8	147309	3	ABK49450	Abk49450 Human tra
c	135	35	77.8	2403	14	ADW98421	Adw98421 Cya codin	C 208	35	77.8	151826	3	AAP22291	Aap22291 BAC conta
c	136	35	77.8	2823	9	ADA02812	Ada02812 Human Fis	C 209	35	77.8	169269	14	ABE35714	Abe35714 L. pneumo
c	137	35	77.8	2823	10	ADB72550	Adb72550 Human CA	C 210	35	77.8	174448	11	ACN43946	Acn43946 Human gen
c	138	35	77.8	2823	10	ADC85292	Adc85292 Human cod	C 211	35	77.8	337344	13	ABD32715	ABd32715 Human can
c	139	35	77.8	2823	12	ADM74407	Adm74407 Human car	C 212	35	77.8	349980	3	AAF21610	Aaf21610 Neisseria
c	140	35	77.8	2990	2	AAQ04123	Aaq04123 Adenyl cy	C 213	34	75.6	375	9	AAL57553	Aal57553 Human zin
c	141	35	77.8	3026	4	AAL62778	Aal62778 Human cDN	C 214	34	75.6	420	3	AAC57697	Aac57697 Arachidon
c	142	35	77.8	3026	4	ABK43822	Abk43822 DNA encod	C 215	34	75.6	446	5	AAS84654	Aas84654 DNA encod
c	143	35	77.8	3026	5	ADM19427	Adm19427 Novel hum	C 216	34	75.6	486	9	ACH28912	Ach28912 Human adu
c	144	35	77.8	3026	12	ADK54209	Adk54209 cDNA enco	C 217	34	75.6	487	4	AAH84032	Aah84032 Sulemar r
c	145	35	77.8	3195	4	AAS22715	Aas22715 Human cDN	C 218	34	75.6	487	4	AAH84035	Aah84035 Sulemar r
c	146	35	77.8	3251	10	ADA53369	Ada53369 Human cod	C 219	34	75.6	488	4	AAH84037	Aah84037 Sulemar r
c	147	35	77.8	3287	3	AAC98900	Aac98900 Human pan	C 220	34	75.6	497	13	ACN59886	Acn59886 Cotton gy
c	148	35	77.8	3437	10	ADB47716	Adb47716 A. gossyp	C 221	34	75.6	503	4	AAF77545	Aaf77545 Human mab
c	149	35	77.8	3529	10	ADI02641	Adi02641 Human cDN	C 222	34	75.6	503	4	AAF77547	Aaf77547 Human mab
c	150	35	77.8	3534	8	ACD13215	AcD13215 cDNA enco	C 223	34	75.6	526	6	ABQ46702	ABq46702 Oligonuc1

c 224	34	75.6	526	6	ABQ46703	Abq46703 Oligonucle	c 297	34	75.6	2334	11	ABD05203	Abd05203 Pseudomon
c 225	34	75.6	531	13	ACN58302	Acn58302 Cotton gy	298	34	75.6	2339	3	AAC50813	Aac50813 Arabidops
c 226	34	75.6	534	4	ABM61910	Abm61910 Human foe	299	34	75.6	2388	3	AAC36244	Aac36244 Arabidops
227	34	75.6	534	4	AAI41833	Aai41833 Probe #10	300	34	75.6	2415	8	ACF72808	Acf72808 Staphyloc
228	34	75.6	534	4	AAK36119	Aak36119 Human bon	301	34	75.6	2418	6	ABN92521	Abn92521 Staphyloc
229	34	75.6	534	4	AAK10220	Aak10220 Human bra	302	34	75.6	2418	13	AD020598	Ado20598 Staphyloc
230	34	75.6	534	4	AB835812	Ab835812 Human liv	303	34	75.6	2503	2	AAX28272	Aax28272 S. aureus
231	34	75.6	534	6	AB810257	Ab810257 Human gen	304	34	75.6	2561	4	AAI98069	Aai98069 Human neu
c 232	34	75.6	555	2	AAK21043	Aak21043 Polynucle	305	34	75.6	2618	13	ACN38162	Acn38162 Tumour-as
c 233	34	75.6	593	6	ABL36763	Ab136763 Human col	306	34	75.6	2648	12	ADO20350	Ado20350 Human PRO
234	34	75.6	601	14	AB3322980	Aeb322980 Human DNA	307	34	75.6	2673	14	ADY18410	Ady18410 DNA encod
235	34	75.6	601	14	AB3322981	Aeb322981 Human DNA	308	34	75.6	2673	4	ABL08213	Ab108213 Drosophil
236	34	75.6	602	3	AA822293	Aaa822293 N. mening	309	34	75.6	2719	6	AB655573	Ab655573 Mouse Vps
c 237	34	75.6	602	6	ABK39526	Abk39526 cDNA enco	310	34	75.6	2745	8	ADA68622	Ada68622 Arabidops
c 238	34	75.6	602	8	ACA11855	Acail1855 Human lun	311	34	75.6	2766	3	AAC49730	Aac49730 Arabidops
c 239	34	75.6	602	8	ACA03041	Acac03041 Lung canc	312	34	75.6	2770	4	AAF77544	Aaf77544 Human mab
c 240	34	75.6	602	10	ADH47083	Adh47083 Human lun	313	34	75.6	2979	3	AAC43074	Aac43074 Arabidops
c 241	34	75.6	602	13	ADJ211002	Adj211002 Human lun	314	34	75.6	3120	8	ACA53888	Aca53888 Prokaryot
c 242	34	75.6	618	10	ADK56055	Adk56055 Plant DNA	315	34	75.6	3370	4	AH54358	Aah54358 S. epider
c 243	34	75.6	626	6	ABK16030	Abk16030 Human lun	316	34	75.6	3421	4	ABL28190	Ab128190 Drosophil
c 244	34	75.6	626	10	ADB95293	Adb95293 Human lun	317	34	75.6	3797	2	AAK28269	Aax28269 S. aureus
245	34	75.6	627	10	ADC08715	Adc08715 Wheat DNA	318	34	75.6	4101	4	ABL03255	Ab103255 Drosophil
246	34	75.6	642	6	ABK78633	Abk78633 Bacillus	319	34	75.6	4384	3	AAC60914	Aac60914 Human squ
c 247	34	75.6	649	8	ABZ18542	Abz18542 Group III	320	34	75.6	4465	6	ABV77992	Abv77992 Hypoxia-r
c 248	34	75.6	717	11	ACH98404	Ach98404 Klebsiell	321	34	75.6	4465	9	ACC57774	Acc57774 Human cyc
249	34	75.6	747	3	AAF14684	Aaf14684 Aspergill	322	34	75.6	4465	10	ACF79931	Acf79931 Breast ca
250	34	75.6	747	13	ADU58725	Adu58725 Aspergill	323	34	75.6	4465	10	ABX08805	Abx08805 Angiogene
251	34	75.6	747	14	AD296728	Adt296728 Aspergill	324	34	75.6	4465	11	ADN38713	Adn38713 Cancer/an
252	34	75.6	750	13	ADT42668	Adt42668 Bacterial	325	34	75.6	4465	11	ADN95605	Adn95605 Human BEC
c 253	34	75.6	836	6	ABZ16106	Abz16106 Arabidops	326	34	75.6	4465	12	ADO24386	Ado24386 Human PRO
254	34	75.6	865	6	AB865571	Ab865571 Mouse Vps	327	34	75.6	4465	12	ADP10451	Adp10451 Reference
255	34	75.6	876	12	ADQ23038	Adq23038 Human sof	328	34	75.6	4465	13	ADQ80275	Adq80275 Prostegla
256	34	75.6	926	6	ABK78104	Abk78104 Bacillus	329	34	75.6	4465	13	ADU05832	Adu05832 Novel bro
257	34	75.6	978	8	ACC85746	Acc85746 Predicted	330	34	75.6	4465	14	ADX85144	Adx85144 Human pro
258	34	75.6	978	8	ACC85747	Acc85747 Experimen	331	34	75.6	4465	14	ADY15397	Ady15397 DNA encod
259	34	75.6	1024	6	ABK65736	Abk65736 Helicobac	332	34	75.6	4465	14	ADY19547	Ady19547 DNA encod
c 260	34	75.6	1026	5	ABV23687	Abv23687 Human pro	333	34	75.6	4465	14	ADZ59960	Adz59960 Human COX
261	34	75.6	1030	5	ABA19089	Abai9089 Human ner	334	34	75.6	4465	14	AEA23664	Aea23664 Human PRO
262	34	75.6	1038	5	ABV28423	Abv28423 Human pro	335	34	75.6	4496	8	ACA03925	Aca03925 cDNA down
263	34	75.6	1030	5	ABV22601	Abv22601 Human pro	336	34	75.6	4496	8	ABX63478	Abx63478 Human cDN
264	34	75.6	1100	2	AA740216	Aat40216 Sequence	337	34	75.6	4602	2	AAV17604	Aav17604 Nucleotid
265	34	75.6	1152	6	AB865572	Ab865572 Mouse Vps	338	34	75.6	4665	14	ADM94116	Adm94116 Staphyloc
266	34	75.6	1152	13	ADT18513	Adt18513 Plant cDN	339	34	75.6	4686	8	ACC84450	Acc84450 Glucanase
267	34	75.6	1165	2	AA772945	Aat72945 Phaffia c	340	34	75.6	4750	10	ADZ52722	Adz52722 Human cDN
268	34	75.6	1224	8	ACA28081	Acas28081 Prokaryot	341	34	75.6	5290	4	ABL08212	Ab108212 Drosophil
c 269	34	75.6	1227	8	ACA30263	Acas30263 Prokaryot	342	34	75.6	5700	11	AE886031	Aeb86031 DNA damag
c 270	34	75.6	1237	5	AA066660	Aad66660 A. thalia	343	34	75.6	5945	6	ABL32085	Ab132085 Human imm
c 271	34	75.6	1237	10	ADB31816	Adb31816 DNA encod	344	34	75.6	6081	6	ABL32287	Ab132287 Human imm
c 272	34	75.6	1237	12	ADO02040	Ado02040 Thalecres	345	34	75.6	6633	14	ADZ13382	Adz13382 Human can
c 273	34	75.6	1268	4	AAV74604	Aav74604 Staphyloc	346	34	75.6	6780	4	AAK69692	Aak69692 Human imm
c 274	34	75.6	1278	4	AAF77543	Aaf77543 Caenorhab	347	34	75.6	6780	4	AAK69694	Aak69694 Human imm
c 275	34	75.6	1326	8	ACA21297	Acas21297 Prokaryot	348	34	75.6	6784	4	AAK69693	Aak69693 Human imm
276	34	75.6	1338	11	ABD05102	Abd05102 Pseudomon	349	34	75.6	6784	4	AAK69691	Aak69691 Human imm
c 277	34	75.6	1352	3	AAK40200	Aac40200 Arabidops	350	34	75.6	6910	4	AAK81265	Aak81265 Human imm
c 278	34	75.6	1368	13	ADX31745	Adx31745 Plant ful	351	34	75.6	6910	4	AAI06121	Aai06121 Human rep
c 279	34	75.6	1437	8	ACA21522	Acas21522 Prokaryot	352	34	75.6	6910	4	ABL98686	Ab198686 Human tes
c 280	34	75.6	1560	9	AD808449	Ad808449 Allostoc	353	34	75.6	8421	10	AB268179	Ab268179 Human sec
281	34	75.6	1560	9	AD808445	Ad808445 Allostoc	354	34	75.6	8506	6	ABQ80961	Abq80961 Dextran s
282	34	75.6	1560	9	AD808447	Ad808447 Allostoc	355	34	75.6	8506	8	ACC70331	Acc70331 Nucleotid
283	34	75.6	1560	9	AD808451	Ad808451 Allostoc	356	34	75.6	8931	6	ABQ80962	Abq80962 Dextran s
284	34	75.6	1560	9	AD808443	Ad808443 Allostoc	357	34	75.6	8931	8	ACC70332	Acc70332 Nucleotid
285	34	75.6	1563	3	ACA41626	Aac41626 Arabidops	358	34	75.6	9451	6	AB897450	Ab897450 Human cyc
c 286	34	75.6	1579	3	AAK81102	Aak81102 Human imm	359	34	75.6	9453	6	AB897450	Ab897450 Human cyc
c 287	34	75.6	1744	3	AAV79724	Aav79724 Pinus rad	360	34	75.6	9453	3	AAK34994	Aak34994 Human ade
288	34	75.6	1759	13	ADT17132	Adt17132 Plant cDN	361	34	75.6	9453	6	ABL65014	Ab165014 Lung canc
289	34	75.6	1781	3	ACA41539	Aac41539 Arabidops	362	34	75.6	9453	6	ABK84194	Abk84194 Human cDN
c 290	34	75.6	1844	4	ABK43462	Abk43462 DNA encod	363	34	75.6	9453	10	AB296810	Ab296810 Human nuc
c 291	34	75.6	1844	12	ADI53849	Adi53849 cDNA enco	364	34	75.6	9453	10	ADK61312	Adk61312 Ovarian c
c 292	34	75.6	1859	6	ABQ33833	Abq33833 Oligonucle	365	34	75.6	9453	11	ABD20659	Abd20659 Human pul
c 293	34	75.6	1859	6	ABQ33832	Abq33832 Oligonucle	366	34	75.6	9453	12	ADP13473	Adp13473 Renal cel
294	34	75.6	2174	13	ADK59745	Adk59745 Plant ful	367	34	75.6	9979	5	AAK80896	Aak80896 Human pro
295	34	75.6	2291	14	ADY37552	Ady37552 Lung canc	368	34	75.6	10316	4	AAK89399	Aak89399 Human dig
296	34	75.6	2301	4	AAH53592	Aah53592 S. epider	369	34	75.6	10589	4	ABL28294	Ab128294 Drosophil

c 370	34	75.6	10997	10	ADL13869	Adl13869 Osteoearth	443	33	73.3	457	6	ABN63885	Abn63885 Human can
c 371	34	75.6	11064	9	ADL20274	Adl20274 Human cyc	444	33	73.3	460	3	ABC31677	Abc31677 Human sec
c 372	34	75.6	11524	8	ABX76362	Abx76362 Lung canc	445	33	73.3	474	6	ABN63815	Abn63815 Human can
c 373	34	75.6	11524	11	ADN39475	Adn39475 Cancer/an	446	33	73.3	483	11	ADM45042	Adm45042 Insect re
c 374	34	75.6	11524	11	ADN38858	Adn38858 Cancer/an	c 447	33	73.3	512	6	ABT09880	Abt09880 Human bre
c 375	34	75.6	14781	4	AAI36303	Aai36303 Human mus	448	33	73.3	514	4	AAI17640	Abt17640 Probe #75
c 376	34	75.6	14781	8	ABX59291	Abx59291 cDNA enco	449	33	73.3	514	4	ABA62577	Abag2577 Human foe
c 377	34	75.6	14781	12	ADJ30041	Adj30041 Human mus	450	33	73.3	514	4	AAI42565	AAI42565 Probe #11
c 378	34	75.6	15240	3	AAA34995	Aaa34995 Human ade	451	33	73.3	514	4	AAK36788	Aak36788 Human bon
c 379	34	75.6	15240	3	AAF21117	Aaf21117 Human low	452	33	73.3	514	4	AAK10934	Aak10934 Human bra
c 380	34	75.6	15240	10	ABZ96811	Abz96811 Human nuc	453	33	73.3	514	6	ABS36451	Abes36451 Human liv
c 381	34	75.6	15240	11	ABD20660	Abd20660 Human pul	454	33	73.3	514	6	ABS10792	Abes10792 Human gen
c 382	34	75.6	16488	6	ADZ25856	Adz25856 Human dip	c 455	33	73.3	549	4	AAH10648	Aah10648 Human CDN
c 383	34	75.6	16488	6	ADZ25895	Adz25895 Human dip	456	33	73.3	551	6	ABV87369	Adv87369 Human col
c 384	34	75.6	21458	4	AAK81266	Aak81266 Human imm	457	33	73.3	561	4	AAF71083	Aaf71083 C. Glutam
c 385	34	75.6	21458	4	AAI06122	Aai06122 Human rep	458	33	73.3	577	4	AAH33041	Aah33041 Human col
c 386	34	75.6	21458	4	ABL98687	Abi98687 Human tes	c 459	33	73.3	578	8	ABZ19157	Abz19157 Group III
c 387	34	75.6	22618	4	ABL03254	Abi03254 Drosophil	460	33	73.3	580	6	ABN65869	Abns65869 Human can
c 388	34	75.6	35331	14	ADW97783	Adw97783 Subgroup	461	33	73.3	580	14	ACL63021	ACL63021 Human col
c 389	34	75.6	37004	13	ABD33353	Abd33353 Human can	c 462	33	73.3	582	6	ABV88033	Adv88033 Human col
c 390	34	75.6	37487	9	ADA02498	Ada02498 Human MYC	463	33	73.3	585	11	ACL32206	Adl32206 Rice ablo
c 391	34	75.6	37487	10	ADB72236	Adb72236 Human MYC	464	33	73.3	585	12	ADJ42266	Adj42266 Plant cDN
c 392	34	75.6	37487	10	ADB82938	Ades82938 Human MYC	465	33	73.3	612	14	ACL62781	ACL62781 Human col
c 393	34	75.6	37487	10	ADB95746	Ades95746 Human MYC	466	33	73.3	618	5	AAH65820	Aah65820 C. Glutam
c 394	34	75.6	42018	14	ADZ13379	Adz13379 Human can	467	33	73.3	618	8	ACA00184	ACA00184 C. Glutam
c 395	34	75.6	46765	6	AA959306	Aas959306 DNA enco	c 468	33	73.3	631	3	AAA75998	Aaa75998 DNA enco
c 396	34	75.6	48775	14	ABE77455	Abew77455 Human TGF	c 469	33	73.3	638	13	ACN54044	ACN54044 Cotton an
c 397	34	75.6	50000	3	AA96364	Aas96364 Polymorph	470	33	73.3	649	4	ABL04891	Aal04891 Human rep
c 398	34	75.6	80264	13	ABD33063	Abd33063 Mouse can	471	33	73.3	649	4	ABL97785	Abi97785 Human tea
c 399	34	75.6	81001	4	AAF30035	Aaf30035 Human apo	c 472	33	73.3	662	14	ADZ60615	Adz60615 Rat g3019
c 400	34	75.6	100848	4	AAF28552	Aaf28552 Genomic f	473	33	73.3	670	10	ADM45312	Adm45312 Plant DNA
c 401	34	75.6	110000	6	ABQ69245	Abq69245 Continuation (12 o	474	33	73.3	670	11	ADM45645	Adm45645 Insect re
c 402	34	75.6	110000	6	ABQ67197	Abq67197 Continuation (11 o	475	33	73.3	696	9	ADA14453	Ada14453 Mouse spe
c 403	34	75.6	110000	9	ADB12064	Abd12064 Continuation (5 o	c 476	33	73.3	732	13	ADR59252	Adr59252 Cotton cD
c 404	34	75.6	110000	9	ADB12064	Abd12064 Continuation (7 o	477	33	73.3	735	5	AAH65659	Aah65659 C. Glutam
c 405	34	75.6	110000	10	AAI52246	Aai52246 Human gen	478	33	73.3	736	4	AAF71084	Aaf71084 C. Glutam
c 406	34	75.6	110000	12	ADQ97328	Adq97328 Continuation (3 of	479	33	73.3	741	4	AAH07435	Aah07435 Human CDN
c 407	34	75.6	110000	13	ABD32923	Abd32923 Continuation (3 of	c 480	33	73.3	774	6	ABQ70300	Abq70300 Listaria
c 408	34	75.6	116592	8	ABX15519	Abx15519 Human tyr	481	33	73.3	788	10	ADD42729	Add42729 Chlamydia
c 409	34	75.6	116592	10	ADA47900	Ada47900 Human tra	482	33	73.3	788	10	ADD42728	Add42728 Chlamydia
c 410	34	75.6	116592	14	ABE47448	Abey47448 Human sul	483	33	73.3	873	2	AAI14851	Aat14851 Potato UI
c 411	34	75.6	120670	12	ADQ59167	Adq59167 MSI-H car	484	33	73.3	873	2	AAI14851	Aat14851 Potato UI
c 412	34	75.6	126990	12	ADP13332	Adp13332 Renal cel	485	33	73.3	882	10	ABX06246	Abx06246 S. pneumo
c 413	34	75.6	173810	6	ABN85752	Abn85752 Mouse chr	486	33	73.3	885	13	ADK43750	Adk43750 Streptoco
c 414	34	75.6	174566	8	ABQ77400	Abq77400 Human ITG	487	33	73.3	895	2	AAI30749	Aai30749 Streptoco
c 415	34	75.6	174566	12	ADL08118	Adl08118 Human gen	c 488	33	73.3	897	3	AAI30749	Aai30749 Streptoco
c 416	34	75.6	175338	11	ACN45088	Acn45088 Mouse gen	489	33	73.3	927	13	ADR91494	Adr91494 Novel S.
c 417	34	75.6	185548	13	ADV34986	Adv34986 Murine cD	490	33	73.3	927	14	AAE55364	Aae55364 Streptoco
c 418	34	75.6	188971	12	ADL08108	Adl08108 Human gen	c 491	33	73.3	939	13	ADS46476	Ads46476 Bacteriat
c 419	34	75.6	207542	14	ABE832385	Abes832385 Human gen	492	33	73.3	1006	3	AAF13852	Aaf13852 Aspergill
c 420	34	75.6	207557	14	ABE832371	Abes832371 Human gen	c 493	33	73.3	1006	13	ADU57893	Adu57893 Aspergill
c 421	34	75.6	220860	12	ADN36595	Adn36595 Human pro	c 494	33	73.3	1006	14	ADZ95896	Adz95896 Aspergill
c 422	34	75.6	254868	14	ADZ13236	Adz13236 Murine ca	495	33	73.3	1017	6	ABQ47300	Abq47300 Oligonuecl
c 423	34	75.6	326002	13	ABD32843	Abd32843 Human can	c 496	33	73.3	1017	6	ABQ47300	Abq47300 Oligonuecl
c 424	33	73.3	115	12	ADQ21828	Adq21828 Human sof	497	33	73.3	1050	4	AAI55163	Aai55163 Staphyloc
c 425	33	73.3	158	3	AAI10763	Aai10763 Human sec	c 498	33	73.3	1059	10	ACF71850	Acf71850 Photorhab
c 426	33	73.3	168	6	ADP36048	Adp36048 Human cHL	c 499	33	73.3	1069	13	ADX11707	Adx11707 Plant ful
c 427	33	73.3	279	10	ADP02537	Adp02537 Bacterial	c 500	33	73.3	1097	13	ADX30260	Adx30260 Plant ful
c 428	33	73.3	308	13	ACN51920	Acn51920 Cotton an	501	33	73.3	1125	8	ACA20181	ACA20181 Prokaryot
c 429	33	73.3	358	6	ABN62313	Abn62313 Human can	502	33	73.3	1125	14	ADW94250	Adw94250 Staphyloc
c 430	33	73.3	387	3	AAE76689	Aae76689 Human ORF	503	33	73.3	1125	14	ADW94599	Adw94599 Prolifera
c 431	33	73.3	391	5	ABV50394	Abv50394 Human pro	504	33	73.3	1134	8	ACF73836	Acf73836 Staphyloc
c 432	33	73.3	395	13	ADP65266	Adp65266 Cotton cD	505	33	73.3	1137	4	AAE54560	Aae54560 Staphyloc
c 433	33	73.3	406	9	ACH29363	Ach29363 Human adu	506	33	73.3	1148	6	AAI33871	Aai33871 Corynebac
c 434	33	73.3	408	2	AAQ060978	Aaq060978 Human bra	c 507	33	73.3	1152	4	AAF61897	Aaf61897 Polyoma v
c 435	33	73.3	408	4	AAI00286	Aai00286 Human rep	c 508	33	73.3	1155	13	ABE19645	Abel19645 Murine po
c 436	33	73.3	409	9	ACH24089	Ach24089 Human adu	509	33	73.3	1178	8	ADA05244	Ada05244 Human sec
c 437	33	73.3	421	8	AAE77542	Aae77542 Brugia ma	c 510	33	73.3	1178	4	ADA40448	Ada40448 Human sec
c 438	33	73.3	421	8	ABX50406	Abx50406 Bovine ES	c 511	33	73.3	1178	9	ADB91397	Adb91397 Human sec
c 439	33	73.3	440	8	ABX50214	Abx50214 Bovine ES	512	33	73.3	1207	12	ADI67114	Adi67114 Novel Lac
c 440	33	73.3	443	8	ABZ35900	Abz35900 Human sec	c 513	33	73.3	1284	3	AAI06342	Aai06342 Human sec
c 441	33	73.3	451	6	ABN64053	Abn64053 Human can	514	33	73.3	1287	6	AAE99128	Aae99128 Human ELA
c 442	33	73.3	452	5	ABV51234	Abv51234 Human pro	515	33	73.3	1347	8	ACA22242	ACA22242 Prokaryot

516	33	73.3	1393	6	ABK36065	Abk36065 cDNA sequ	c 589	73.3	3794	6	ABK24531	Abk24531 EIP-2alph
517	33	73.3	1457	12	AD019686	Ado19686 Human PRO	c 590	73.3	3832	6	ABN59604	Abn59604 Novel hum
518	33	73.3	1517	12	ADM18451	Adm18451 Human chr	c 591	73.3	3951	6	ABA93739	Aba93739 Human int
519	33	73.3	1565	3	ACAC77295	Aac77295 Human ORF	c 592	73.3	3956	8	ABT19522	Abt19522 Aspergill
520	33	73.3	1605	12	AD015981	Ado15981 4 synthet	c 593	73.3	4027	10	AAD49463	Aad49463 Human ves
521	33	73.3	1628	10	ADB56180	Adb56180 Toxicity-	c 594	73.3	4104	2	AAK07356	Aak07356 Arabidops
522	33	73.3	1628	10	ADB50703	Adb50703 Primary r	c 595	73.3	4104	9	ACA62105	Ac62105 cDNA enco
523	33	73.3	1628	11	ADW22335	Adw22335 Rat hepat	c 596	73.3	4308	13	ADR07363	Adr07363 Pull leng
524	33	73.3	1654	12	AD167021	Adi167021 Novel Lac	c 597	73.3	4635	13	ADR84751	Adr84751 Aspergill
525	33	73.3	1684	11	ACN91123	Acn91123 Breast ca	c 598	73.3	4700	5	AAF54982	Aaf54982 Nucleotid
526	33	73.3	1767	8	ACA28630	Aca28630 Prokaryot	c 599	73.3	4700	5	AAD30800	Aad30800 Braeasin
527	33	73.3	1800	6	ABZ13816	Abz13816 Arabidops	c 600	73.3	4707	5	AAS88600	Aas88600 DNA enco
528	33	73.3	1800	8	ABZ42024	Abz42024 Arabidops	c 601	73.3	4820	12	ADQ63312	Adq63312 Novel hum
529	33	73.3	1800	8	ADA68334	Ada68334 Arabidops	c 602	73.3	4859	11	ACN88670	Acn88670 Breast ca
530	33	73.3	1853	4	AAH14804	Aah14804 Human cDN	c 603	73.3	4906	4	AAL26684	Aal26684 Human bre
531	33	73.3	1874	13	ADU51489	Adu51489 Potato et	c 604	73.3	4989	5	ABA16291	Ab16291 Human ner
532	33	73.3	1923	2	AAK511745	Aak511745 DNA enco	c 605	73.3	4989	5	ABA18318	Ab18318 Human ner
533	33	73.3	1923	6	ABQ92597	Abq92597 Human sec	c 606	73.3	5296	10	ADD25527	Add25527 Blding d
534	33	73.3	2139	3	AAZ29197	Aaz29197 Human mye	c 607	73.3	5297	3	AAA62084	Aaa62084 Polyoma T
535	33	73.3	2139	3	AAA50667	Aaa50667 DNA enco	c 608	73.3	5310	13	ADR06767	Adr06767 Full leng
536	33	73.3	2139	6	ABK88940	Abk88940 Human mye	c 609	73.3	5373	14	ADX17031	Adx17031 Lemna min
537	33	73.3	2139	6	ABN86590	Abn86590 Human MBP	c 610	73.3	5393	6	ABT07728	Abt07728 Breast ca
538	33	73.3	2139	14	ARB94281	Aeb94281 Human mye	c 611	73.3	5436	10	ADI40468	Adi40468 Human pur
539	33	73.3	2145	4	AAI59701	Aai59701 Human pol	c 612	73.3	5809	10	ADI02562	Adi02562 Human cDN
540	33	73.3	2156	4	AAK98996	Aak98996 Human pan	c 613	73.3	5852	6	ABT13872	Abt13872 Human hel
541	33	73.3	2165	4	ABL06091	Ab106091 Drosophil	c 614	73.3	5852	11	ADN60077	Adn60077 Human hel
542	33	73.3	2165	13	ADV35074	Adv35074 Human cDN	c 615	73.3	5872	4	ABL06090	Ab106090 Drosophil
543	33	73.3	2187	14	ACL73120	Ac173120 M. xanthu	c 616	73.3	5937	6	ABL34543	Ab134543 Human met
544	33	73.3	2199	12	ADQ22766	Adq22766 Human eof	c 617	73.3	5937	7	ADN80151	Adn80151 Human che
545	33	73.3	2204	10	ADE58438	Ade58438 Human gen	c 618	73.3	5937	7	ADN99804	Adn99804 Complemen
546	33	73.3	2204	10	ADE58434	Ade58434 Human gen	c 619	73.3	6036	6	ABL33108	Ab133108 Human imm
547	33	73.3	2220	10	ACF70571	Acf70571 Photorhab	c 620	73.3	6036	6	ABK31290	Abk31290 Signal tr
548	33	73.3	2222	14	ADW18170	Adw18170 Pinus rad	c 621	73.3	6036	6	ABL70267	Ab170267 Chemicall
549	33	73.3	2244	2	AAZ96309	Aaz96309 S. pneumo	c 622	73.3	6036	6	AAK61192	Aak61192 Human gen
550	33	73.3	2245	12	ADQ86500	Adq86500 Human tum	c 623	73.3	6305	4	AAK06437	Aak06437 Human rep
551	33	73.3	2279	4	AAH15843	Aah15843 Human cDN	c 624	73.3	6305	5	AAK40549	Aak40549 DNA enco
552	33	73.3	2279	5	AAI93873	Aai93873 Human sto	c 625	73.3	6305	11	ADJ09755	Adj09755 Human pro
553	33	73.3	2295	13	ADR85925	Adr85925 Aspergill	c 626	73.3	6343	3	AAA57891	Aaa57891 Maize glo
554	33	73.3	2310	11	ACH99235	Ach99235 Klebsiell	c 627	73.3	6456	6	ABL33006	Ab133006 Human imm
555	33	73.3	2319	10	ADB79793	Adb79793 Human put	c 628	73.3	6495	4	AAK06441	Aak06441 Human rep
556	33	73.3	2336	4	AAK86408	Aak86408 Human imm	c 629	73.3	6495	5	AAK40553	Aak40553 DNA enco
557	33	73.3	2336	4	AAK68616	Aak68616 Human imm	c 630	73.3	6495	11	ADJ09759	Adj09759 Human pro
558	33	73.3	2358	11	ADM01560	Adm01560 Human cDN	c 631	73.3	6750	5	ABL08321	Ab108321 Drosophil
559	33	73.3	2403	4	AAI61129	Aai61129 Human pol	c 632	73.3	6750	5	AAI14363	Aai14363 Drosophil
560	33	73.3	2458	4	AAI59343	Aai59343 Human pol	c 633	73.3	6820	4	ABL12056	Ab112056 Drosophil
561	33	73.3	2494	11	AAU88758	Aau88758 Mouse Sha	c 634	73.3	6839	4	ABL15964	Ab115964 Drosophil
562	33	73.3	2496	2	AAV26572	Aav26572 Carboxydo	c 635	73.3	6941	10	ADC30320	Adc30320 Human nov
563	33	73.3	2521	13	ACN39830	Acn39830 Tumour-as	c 636	73.3	6996	12	ADO19684	Ado19684 Human PRO
564	33	73.3	2537	4	ABL10506	Ab110506 Drosophil	c 637	73.3	6996	13	ADP54810	Adp54810 Human PRO
565	33	73.3	2538	14	ADZ49769	Adz49769 Insulin s	c 638	73.3	6996	13	ADP24218	Adp24218 PRO polyp
566	33	73.3	2587	2	AAI15761	Aai15761 Gibberell	c 639	73.3	7037	6	ABT13865	Abt13865 Human hel
567	33	73.3	2635	13	ADR85338	Adr85338 Aspergill	c 640	73.3	7037	8	ACD13378	Adc13378 Human DNA
568	33	73.3	2655	13	ADS47456	Ads47456 Bacterial	c 641	73.3	7037	11	ADN60070	Adn60070 DNA enco
569	33	73.3	2723	12	ADQ63162	Adq63162 Novel hum	c 642	73.3	7324	13	ADT07436	Adt07436 Human col
570	33	73.3	2740	12	ADQ67478	Adq67478 Novel hum	c 643	73.3	7345	13	ADT07437	Adt07437 Human col
571	33	73.3	2816	10	ADL13915	Adl13915 Osteoarth	c 644	73.3	7421	8	ADA41598	Ada41598 Human sec
572	33	73.3	2816	12	ADP27372	Adp27372 Human epo	c 645	73.3	7421	9	ADB91890	Adb91890 Human sec
573	33	73.3	2816	12	ADQ18216	Adq18216 Human eof	c 646	73.3	7750	4	ABL04664	Ab104664 Drosophil
574	33	73.3	2852	11	ACN92842	Acn92842 Breast ca	c 647	73.3	7970	4	AAS27697	Aas27697 DNA enco
575	33	73.3	2856	5	AAK87447	Aak87447 DNA enco	c 648	73.3	7970	10	AAK94500	Aak94500 Novel hum
576	33	73.3	2883	11	ACL26239	ACL26239 Rice abio	c 649	73.3	8040	4	AAK05543	Aak05543 Human rep
577	33	73.3	2883	12	ADJ39527	Adj39527 Plant cDN	c 650	73.3	8835	4	ABL02806	Ab102806 Human rep
578	33	73.3	2978	3	AAA96705	Aaa96705 Reporter	c 651	73.3	8922	4	ABL08320	Ab108320 Drosophil
579	33	73.3	3002	11	ADZ51821	Adz51821 Saccharom	c 652	73.3	9717	4	ABK42694	Abk42694 Genomic s
580	33	73.3	3027	13	ADN99087	Adn99087 Dcr-1 hom	c 653	73.3	9717	4	ABK03295	Abk03295 Human rep
581	33	73.3	3156	11	ADM01852	Adm01852 Human cDN	c 654	73.3	9717	9	ADN60850	Adn60850 Connectiv
582	33	73.3	3208	12	ADI61743	Adi61743 Human cDN	c 655	73.3	9741	12	ADQ59183	Adq59183 MSI-H car
583	33	73.3	3208	14	AEA43908	Aea43908 Human cDN	c 656	73.3	9834	2	AAV74348	Aav74348 Staphyloc
584	33	73.3	3327	8	ABT17708	Abt17708 Aspergill	c 657	73.3	10220	12	ADJ62778	Adj62778 Human cDN
585	33	73.3	3327	11	ACN90475	Acn90475 Breast ca	c 658	73.3	10220	12	ADL97801	Adl97801 Human dcr
586	33	73.3	3491	12	ADQ63266	Adq63266 Novel hum	c 659	73.3	10336	5	ABA21257	Aba21257 Human ner
587	33	73.3	3591	6	AB213376	Ab213376 Arabidops	c 660	73.3	10427	12	ADJ81645	Adj81645 Human tyr
588	33	73.3	3722	4	ABL12057	Ab112057 Drosophil	c 661	73.3	10663	5	ABA18961	Ab18961 Human ner

c 662	33	73.3	11425	4	AAK79581	Aak79581 Human imm	c 735	33	73.3	110000	12	ADN46464_14	Continuation (15 o
c 663	33	73.3	11566	14	ACL64625	ACL64625 M. xanthu	c 736	33	73.3	110000	12	ADN47960_06	Continuation (7 of
c 664	33	73.3	12038	4	AAO06668	AAO06668 Human fsh	c 737	33	73.3	110000	12	ADQ97138_1	Continuation (2 of
c 665	33	73.3	13123	4	AAO04954	AAO04954 Human rep	c 738	33	73.3	110000	14	ADZ12814_3	Continuation (4 of
c 666	33	73.3	13123	4	ABL97848	ABL97848 Human tes	c 739	33	73.3	110000	14	ADZ13631_0	Ad13631 Human can
c 667	33	73.3	13131	5	ABA19154	ABA19154 Human ner	c 740	33	73.3	110000	14	ADZ13631_1	Continuation (2 of
c 668	33	73.3	13131	5	ABA16290	ABA16290 Human ner	c 741	33	73.3	110000	14	ADZ13631_2	Continuation (3 of
c 669	33	73.3	13131	5	ABA18317	ABA18317 Human ner	c 742	33	73.3	110000	14	ADZ13620_0	Ad13620 Human can
c 670	33	73.3	13712	8	ADA41636	ADA41636 Human sec	c 743	33	73.3	110000	14	ADZ13620_1	Continuation (2 of
c 671	33	73.3	13712	10	ADA57768	Ada57768 BAC fragm	c 744	33	73.3	110000	14	ADZ13620_2	Continuation (3 of
c 672	33	73.3	19521	4	AAK81193	Aak81193 Human imm	c 745	33	73.3	110000	14	ADZ42274_2	Continuation (6 of
c 673	33	73.3	19521	8	ADA1637	Ada1637 Human sec	c 746	33	73.3	110000	14	ABE39175_05	Continuation (6 of
c 674	33	73.3	19521	10	ADA57769	Ada57769 BAC fragm	c 747	33	73.3	110000	14	ABE42401_05	Continuation (6 of
c 675	33	73.3	19976	4	ABLO3988	Ab103988 Drosophili	c 748	33	73.3	110300	13	ADS36499_0	Ad36499 Human aut
c 676	33	73.3	20365	4	AAK82631	Aak82631 Human imm	c 749	33	73.3	117750	13	ABD33653	Abd33653 Human can
c 677	33	73.3	20365	4	AAK70165	Aak70165 Human imm	c 750	33	73.3	121062	12	ADQ97313	Adq97313 Human can
c 678	33	73.3	23579	10	ADC87112	Adc87112 Human GPC	c 751	33	73.3	122779	12	ADQ97053	Adq97053 Mouse can
c 679	33	73.3	25975	4	AAO04955	Aa104955 Human rep	c 752	33	73.3	144792	10	ADC87620	Adc87620 Human GPC
c 680	33	73.3	25975	4	AAO04955	Aa104955 Human rep	c 753	33	73.3	160271	4	AAF85116	Aaf85116 Nucleotid
c 681	33	73.3	28882	2	AAV52273	Aav52273 Streptoco	c 754	33	73.3	160271	4	AAF85156	Aaf85156 Human chr
c 682	33	73.3	35524	2	AAV22140	Aav22140 Chimpanze	c 755	33	73.3	160271	4	AAF85750	Aaf85750 Bipolar a
c 683	33	73.3	38605	11	ACN44050	Acn44050 Human gen	c 756	33	73.3	160271	4	AAO06667	Aao06667 Human chr
c 684	33	73.3	41765	4	AAK76675	Aak76675 Human imm	c 757	33	73.3	160271	4	AAO04864	Aao04864 Human chr
c 685	33	73.3	41772	4	AAK76676	Aak76676 Human imm	c 758	33	73.3	160271	5	AAH23764	Aah23764 Human chr
c 686	33	73.3	43039	14	ADZ13665_4	Continuation (5 of	c 759	33	73.3	160271	5	AAH40397	Aah40397 150kb fra
c 687	33	73.3	45736	13	ABD33564	Abd33564 Murine ca	c 760	33	73.3	160271	5	AAO04858	Aao04858 Human chr
c 688	33	73.3	55829	13	ABD33512	Abd33512 Human can	c 761	33	73.3	160820	8	ABQ76673	Abq76673 Androgen
c 689	33	73.3	57243	11	ACN44826	Acn44826 Human gen	c 762	33	73.3	186591	8	ACF62750	Acf62750 Cancer ba
c 690	33	73.3	58215	11	ACN44100	Acn44100 Mouse gen	c 763	33	73.3	186591	8	ADB20869	Adb20869 MRP1 base
c 691	33	73.3	58708	4	AAK64739	Aak64739 Human imm	c 764	33	73.3	186591	10	ADB87958	Adb87958 Human UGT
c 692	33	73.3	59554	9	ADA02696	Ada02696 Human TK2	c 765	33	73.3	186591	10	ADB96941	Adb96941 Human MDR
c 693	33	73.3	59554	10	ADB72434	Adb72434 Human TK2	c 766	33	73.3	186591	10	ADB92132	Adb92132 Human MDR
c 694	33	73.3	59554	10	ADE95944	Ade95944 Human TK2	c 767	33	73.3	197997	10	AAI54074	Aai54074 Human tra
c 695	33	73.3	63000	6	ABS67634	Abs67634 Human cas	c 768	33	73.3	208648	8	ACF62735	Acf62735 Cancer ba
c 696	33	73.3	63266	12	ADO48536	Ado48536 Human neu	c 769	33	73.3	208648	8	ACF62740	Acf62740 Cancer ba
c 697	33	73.3	63266	14	ADX80721	Adx80721 Human nid	c 770	33	73.3	208648	8	ADB20850	Adb20850 MRP1 base
c 698	33	73.3	68355	8	ACF62737	ACF62737 Cancer ba	c 771	33	73.3	208648	8	ADB20855	Adb20855 MRP1 base
c 699	33	73.3	68355	8	ADB20852	Adb20852 MRP1 base	c 772	33	73.3	208648	10	ADB87944	Adb87944 Human UGT
c 700	33	73.3	68355	10	ADB87941	Adb87941 Human UGT	c 773	33	73.3	208648	10	ADB87939	Adb87939 Human UGT
c 701	33	73.3	68355	10	ADB96924	Adb96924 Human MDR	c 774	33	73.3	208648	10	ADB96922	Adb96922 Human MDR
c 702	33	73.3	68355	10	ADB92115	Adb92115 Human MDR	c 775	33	73.3	208648	10	ADB96927	Adb96927 Human MDR
c 703	33	73.3	69000	14	ADZ42274_3	Continuation (4 of	c 776	33	73.3	208648	10	ADB92113	Adb92113 Human MDR
c 704	33	73.3	69652	13	ABD33115	Abd33115 Human can	c 777	33	73.3	208648	10	ADB92118	Adb92118 Human MDR
c 705	33	73.3	73038	12	ADQ59401	Adq59401 Human can	c 778	33	73.3	243335	14	ABE42735	Aeb42735 L. pneumo
c 706	33	73.3	73038	14	ADZ13670	Adz13670 Murine ca	c 779	33	73.3	247544	12	ADQ59419	Adq59419 Human can
c 707	33	73.3	76080	14	ADZ12902	Adz12902 Murine ca	c 780	33	73.3	247544	12	ADZ13712	Adz13712 Murine ca
c 708	33	73.3	79084	12	ADQ97563	Adq97563 Murine ca	c 781	33	73.3	256294	13	ABD33020	Abd33020 Mouse can
c 709	33	73.3	80000	12	ADP49338	Adp49338 Human B-c	c 782	33	73.3	295096	11	ACN44068	Acn44068 Mouse gen
c 710	33	73.3	81656	12	ADQ97876	Adq97876 Human can	c 783	33	73.3	295644	14	ABE35721	Aeb35721 L. pneumo
c 711	33	73.3	94400	12	ADP08387	Adp08387 Human gly	c 784	33	73.3	298667	14	ABE39173	Aeb39173 L. pneumo
c 712	33	73.3	110000	2	AAV21209_02	Continuation (3 of	c 785	33	73.3	298667	12	ADQ59380	Adq59380 Human can
c 713	33	73.3	110000	2	AAV21209_03	Continuation (4 of	c 786	33	73.3	310122	13	ABD32533	Abd32533 Mouse can
c 714	33	73.3	110000	2	AAV21209_04	Continuation (5 of	c 787	33	73.3	310122	14	ADZ13032	Adz13032 Murine ca
c 715	33	73.3	110000	2	AAV21209_09	Continuation (10 o	c 788	33	73.3	319608	3	AAH51601	Aah51601 Human chr
c 716	33	73.3	110000	2	AAZ01425_07	Continuation (8 of	c 789	33	73.3	319608	5	AAH509301	Aas09301 Human sch
c 717	33	73.3	110000	6	ABO67196_2	Continuation (3 of	c 790	33	73.3	349980	5	AAH68526	Aah68526 C glutami
c 718	33	73.3	110000	6	ABO69245_00	Abg69245 Listeria	c 791	33	73.3	349980	5	AAH68529	Aah68529 C glutami
c 719	33	73.3	110000	6	ABO69245_05	Continuation (6 of	c 792	33	73.3	349980	5	AAH68528	Aah68528 C glutami
c 720	33	73.3	110000	6	ABO67197_04	Continuation (5 of	c 793	32	71.1	148	2	AAH86204	Aah86204 Human sin
c 721	33	73.3	110000	6	ABO3041_00	Abao3041 Listeria	c 794	32	71.1	151	3	AAA45749	Aaa45749 Human sec
c 722	33	73.3	110000	6	ABO3041_01	Continuation (2 of	c 795	32	71.1	151	3	AAA45749	Aaa45749 Human sec
c 723	33	73.3	110000	6	ABO3041_12	Continuation (13 o	c 796	32	71.1	201	13	ADQ41808	Adq41808 Myocardia
c 724	33	73.3	110000	10	ABE56454_04	Continuation (5 of	c 797	32	71.1	201	13	ADQ41807	Adq41807 Myocardia
c 725	33	73.3	110000	10	ABE56454_36	Continuation (37 o	c 798	32	71.1	201	13	ADQ41712	Adq41712 Myocardia
c 726	33	73.3	110000	10	ACF67367_50	Continuation (51 o	c 799	32	71.1	201	13	ADQ41679	Adq41679 Myocardia
c 727	33	73.3	110000	10	ACF65387_2	Continuation (3 of	c 800	32	71.1	201	13	ADQ41659	Adq41659 Myocardia
c 728	33	73.3	110000	10	ACF65388_11	Continuation (12 o	c 801	32	71.1	201	13	ADQ41634	Adq41634 Myocardia
c 729	33	73.3	110000	10	AAI52246_2	Continuation (3 of	c 802	32	71.1	201	13	ADQ41732	Adq41732 Myocardia
c 730	33	73.3	110000	10	ABQ84281_1	Continuation (2 of	c 803	32	71.1	201	13	ADQ41779	Adq41779 Myocardia
c 731	33	73.3	110000	12	ADN46845_14	Continuation (15 o	c 804	32	71.1	201	13	ADQ41587	Adq41587 Myocardia
c 732	33	73.3	110000	12	ADN47591_06	Continuation (7 of	c 805	32	71.1	201	13	ADQ41660	Adq41660 Myocardia
c 733	33	73.3	110000	12	ADN46123_14	Continuation (15 o	c 806	32	71.1	201	13	ADQ41711	Adq41711 Myocardia
c 734	33	73.3	110000	12	ADN47209_06	Continuation (7 of	c 807	32	71.1	201	13	ADS35173	Ada35173 Human aut

808	32	71.1	201	13	ADS35140	Ad3515140 Human aut	C 881	32	71.1	401	4	AAK96368	Aak96368 Human neu
809	32	71.1	201	13	ADS35166	Ad3515166 Human aut	C 882	32	71.1	401	4	AAK97861	Aak97861 Human neu
810	32	71.1	201	13	ADS35199	Ad3515199 Human aut	C 883	32	71.1	401	6	ABT01138	Abt01138 Human neu
811	32	71.1	201	13	ADS35194	Ad3515194 Human aut	C 884	32	71.1	401	6	AAI02631	Aai02631 Human neu
812	32	71.1	201	13	ADS35171	Ad3515171 Human aut	C 885	32	71.1	410	4	AAI82337	Aai82337 Human pol
813	32	71.1	201	13	ADS37785	Ad37785 Human aut	C 886	32	71.1	414	5	ABV38428	Abv38428 Human pro
814	32	71.1	201	13	ADS35107	Ad35107 Human aut	C 887	32	71.1	414	5	ABV08526	Abv08526 Human pro
815	32	71.1	201	13	ADS35086	Ad35107 Human aut	C 888	32	71.1	418	4	AAI81426	Aai81426 Human pol
816	32	71.1	201	13	ADS35096	Ad35086 Human aut	C 888	32	71.1	426	5	ABV31048	Abv31048 Human pro
817	32	71.1	201	13	ADS35209	Ad35209 Human aut	C 889	32	71.1	430	12	ADK67929	Adk67929 Mouse cdn
818	32	71.1	201	13	ADS35215	Ad35215 Human aut	C 891	32	71.1	435	5	ABV53242	Abv53242 Human pro
819	32	71.1	201	13	ADS35181	Ad35181 Human aut	C 892	32	71.1	435	10	ABX86419	Abx86419 Corn ear-
820	32	71.1	201	13	ADS35225	Ad35225 Human aut	C 893	32	71.1	442	12	ADP95128	Adp95128 Cotton ex
821	32	71.1	201	13	ADS37749	Ad37749 Human aut	C 894	32	71.1	444	12	ADKI6792	Adki6792 Nanaoarcha
822	32	71.1	201	13	ADS35084	Ad35084 Human aut	C 895	32	71.1	444	13	ADUI1910	Adui1910 Solid tum
823	32	71.1	201	13	ADS35200	Ad35200 Human aut	C 896	32	71.1	456	12	ADQ18571	Adq18571 Human sof
824	32	71.1	201	13	ADS37824	Ad37824 Human aut	C 897	32	71.1	456	13	ADQ18571	Adq18571 Novel can
825	32	71.1	201	13	ADS35236	Ad35236 Human aut	C 898	32	71.1	461	5	ABV49438	Abv49438 Human pro
826	32	71.1	201	13	ADS35097	Ad35097 Human aut	C 899	32	71.1	462	6	ABK30377	Abk30377 Human G-p
827	32	71.1	201	13	ADS35214	Ad35214 Human aut	C 900	32	71.1	466	13	ACF90804	Acf90804 Human SIR
828	32	71.1	201	13	ADS35169	Ad35169 Human aut	C 901	32	71.1	468	9	ACH24493	Ach24493 Human adu
829	32	71.1	201	13	ADS35210	Ad35210 Human aut	C 902	32	71.1	471	4	AAH52914	Aah52914 S. epider
830	32	71.1	201	13	ADS35087	Ad35087 Human aut	C 903	32	71.1	472	3	AAC07800	Aac07800 Human sec
831	32	71.1	201	13	ADS35136	Ad35136 Human aut	C 904	32	71.1	472	6	ABN26641	Abn26641 Human ORP
832	32	71.1	201	13	ADS35168	Ad35168 Human aut	C 905	32	71.1	478	6	ABL81251	AbL81251 Human ova
833	32	71.1	201	13	ADS35169	Ad35169 Human aut	C 906	32	71.1	484	5	ACH48342	Ach48342 Human lun
834	32	71.1	201	13	ADS35216	Ad35216 Human aut	C 907	32	71.1	486	5	ABV20476	Abv20476 Human pro
835	32	71.1	201	13	ADS37605	Ad37605 Human aut	C 908	32	71.1	486	5	ABV26314	Abv26314 Human pro
836	32	71.1	201	13	ADS37823	Ad37823 Human aut	C 909	32	71.1	486	5	ABV59627	Abv59627 Human pro
837	32	71.1	201	13	ADS35071	Ad35071 Human aut	C 910	32	71.1	486	5	ABV26478	Abv26478 Human pro
838	32	71.1	201	13	ADS35223	Ad35223 Human aut	C 911	32	71.1	486	5	ABV20637	Abv20637 Human pro
839	32	71.1	201	13	ADS35224	Ad35224 Human aut	C 912	32	71.1	490	4	ABA46411	AbA46411 Human bre
840	32	71.1	201	13	ADS35263	Ad35263 Human aut	C 913	32	71.1	490	6	ABSO5304	Abso5304 Human gen
841	32	71.1	201	13	ADS35151	Ad35151 Human aut	C 914	32	71.1	491	7	ADS71577	AdS71577 Human kid
842	32	71.1	225	6	ABN78300	Abn78300 Human ORF	C 915	32	71.1	491	7	ADS72051	AdS72051 Human kid
843	32	71.1	233	6	ABK45907	Abk45907 cDNA enco	C 916	32	71.1	491	7	ADM40905	Adm40905 cDNA elev
844	32	71.1	245	12	ADL83720	AdL83720 DNA up-re	C 917	32	71.1	491	7	ADM40431	Adm40431 cDNA elev
845	32	71.1	245	12	ADL83721	AdL83721 DNA up-re	C 918	32	71.1	492	6	ABN91831	Abn91831 Staphyloc
846	32	71.1	263	6	ABL37716	AbL37716 Human col	C 919	32	71.1	492	13	ADS02368	AdS02368 Staphyloc
847	32	71.1	287	6	ABL67810	AbL67810 Oesophagu	C 920	32	71.1	494	6	ABI19945	Abi19945 Mouse isc
848	32	71.1	287	6	ABL63043	AbL63043 Breast ca	C 921	32	71.1	497	3	AAC40997	Aac40997 Zee may8
849	32	71.1	287	6	ABL63261	AbL63261 Breast ca	C 922	32	71.1	510	5	AAF68086	Aaf68086 Human lun
850	32	71.1	287	6	ABN93559	Abn93559 Gene #57	C 923	32	71.1	510	6	ABK37997	Abk37997 cDNA enco
851	32	71.1	299	6	ABL73685	AbL73685 Corn tass	C 924	32	71.1	510	8	ACA10326	AcA10326 Human lun
852	32	71.1	327	2	AAV20164	Aav20164 Probe (65	C 925	32	71.1	510	8	ABX99277	Abx99277 Lung canc
853	32	71.1	329	5	ABV19671	Abv19671 Human pro	C 926	32	71.1	510	10	ADH45523	Adh45523 Human lun
854	32	71.1	336	6	ABL78150	AbL78150 Human ova	C 927	32	71.1	510	12	ADP72060	Adp72060 Human lun
855	32	71.1	337	6	ABN17262	Abn17262 Human ORF	C 928	32	71.1	510	13	ADJ19442	Adj19442 Human lun
856	32	71.1	344	9	ACH31266	Ach31266 Human bon	C 929	32	71.1	511	6	ABN24510	Abn24510 Human ORF
857	32	71.1	345	11	ACH97832	Ach97832 Klebeieil	C 930	32	71.1	513	12	ACH73823	Ach73823 Human gen
858	32	71.1	349	5	ADI73416	Adi73416 Human ova	C 931	32	71.1	515	14	ADV75886	Adv75886 Human col
859	32	71.1	349	5	ADL38546	AdL38546 Human ova	C 932	32	71.1	516	5	ABV40016	Abv40016 Human pro
860	32	71.1	354	10	ADK61417	AdK61417 Ovarian c	C 933	32	71.1	516	13	ADS55010	AdS55010 Bacterial
861	32	71.1	367	4	AAI14495	Aai14495 Probe #44	C 934	32	71.1	521	13	ADUI0581	Adui0581 Solid tum
862	32	71.1	367	4	ABA56224	AbA56224 Human foe	C 935	32	71.1	524	2	AAZ41994	Aaz41994 Human end
863	32	71.1	367	4	AAI35871	Aai35871 Probe #45	C 936	32	71.1	526	12	ADQ21758	Adq21758 Human sof
864	32	71.1	367	4	ABA45718	AbA45718 Human bre	C 937	32	71.1	526	13	ACF83913	AcF83913 Human SIR
865	32	71.1	367	4	ABA25869	AbA25869 Probe #43	C 938	32	71.1	536	13	ACN49585	Acn49585 Cotton pr
866	32	71.1	367	4	AAK29908	Aak29908 Human bon	C 939	32	71.1	537	6	ABT10064	Abt10064 Human bre
867	32	71.1	367	4	AAK04412	Aak04412 Human bra	C 940	32	71.1	537	10	ADH28796	Adh28796 Human chr
868	32	71.1	367	4	ABS29552	Abs29552 Human liv	C 941	32	71.1	537	13	ADT90441	Adt90441 Human gen
869	32	71.1	367	5	AAI04322	Aai04322 Probe #43	C 942	32	71.1	541	13	ACN55707	Acn55707 Cotton an
870	32	71.1	367	6	ABS04474	AbS04474 Human gen	C 943	32	71.1	545	9	ACH35429	Ach35429 Human end
871	32	71.1	375	4	AAI189797	Aai189797 Human pol	C 944	32	71.1	550	4	AAS42989	Aas42989 DNA enco
872	32	71.1	375	5	ADP93504	Adp93504 cDNA enco	C 945	32	71.1	556	13	ACN61815	Acn61815 Cotton gy
873	32	71.1	378	12	ADJ44910	AdJ44910 Plant cdn	C 946	32	71.1	564	9	ACH28751	Ach28751 Human adu
874	32	71.1	379	10	ABT22736	Abt22736 Breast ca	C 947	32	71.1	574	13	ACN46663	Acn46663 Cotton pr
875	32	71.1	381	6	ABS61418	AbS61418 Prostate	C 948	32	71.1	584	4	AAI16809	Aai16809 Probe #67
876	32	71.1	381	10	ABT22771	Abt22771 Breast ca	C 949	32	71.1	584	4	ABA60522	AbA60522 Human foe
877	32	71.1	389	6	ABL78260	AbL78260 Human ova	C 950	32	71.1	584	4	AAI40411	Aai40411 Probe #90
878	32	71.1	396	3	ADF57168	Adf57168 Urogenita	C 951	32	71.1	584	4	ABA28691	AbA28691 Probe #71
879	32	71.1	397	5	ABV09874	Abv09874 Human pro	C 952	32	71.1	584	4	AAK34690	Aak34690 Human bon
880	32	71.1	400	5	ABV00705	Abv00705 Human pro	C 953	32	71.1	584	4	AAK08803	Aak08803 Human bra

```

954 32 71.1 584 4 ABS34463
955 32 71.1 584 6 ABS09283
956 32 71.1 596 6 ABQ20363
957 32 71.1 596 6 ABQ20362
958 32 71.1 596 6 ABQ57307
959 32 71.1 606 13 ACN56475
960 32 71.1 613 12 ADJ40364
961 32 71.1 617 5 ABV59145
962 32 71.1 618 13 ADS48153
963 32 71.1 626 6 ABQ43783
964 32 71.1 626 6 ABQ43782
965 32 71.1 636 5 ABV12580
966 32 71.1 651 12 ADK16832
967 32 71.1 666 10 ADC77269
968 32 71.1 666 10 ADK59933
969 32 71.1 666 10 ADK57152
970 32 71.1 666 10 ADK57326
971 32 71.1 691 6 ABT09278
972 32 71.1 691 10 ADG31012
973 32 71.1 691 12 ADG45647
974 32 71.1 691 12 ADH22935
975 32 71.1 692 2 ADR02380
976 32 71.1 697 5 ADL63772
977 32 71.1 715 5 ABV25474
978 32 71.1 717 2 ADR02073
979 32 71.1 736 4 AA195122
980 32 71.1 739 4 AA195567
981 32 71.1 792 6 ABQ26378
982 32 71.1 792 6 ABQ26379
983 32 71.1 792 11 ACH97769
984 32 71.1 832 4 AAH04768
985 32 71.1 834 8 ACD26286
986 32 71.1 836 6 ABZ15457
987 32 71.1 844 12 ADQ22980
988 32 71.1 862 10 ADK58242
989 32 71.1 862 10 ADK54149
990 32 71.1 888 6 ABQ66969
991 32 71.1 910 12 ADQ25555
992 32 71.1 960 9 ADA30629
993 32 71.1 977 13 ADS61127
994 32 71.1 978 3 AAC39924
995 32 71.1 978 8 ACA26583
996 32 71.1 1004 3 AAZ97180
997 32 71.1 1006 4 ABL14243
998 32 71.1 1007 13 ADT05028
999 32 71.1 1008 13 ADS49042
1000 32 71.1 1017 8 ACA48331

Breast cancer differentially expressed gene product #47.
ds: cytostatic; gene therapy; DKFZp5661133 activity inhibitor;
breast cancer; differential expression.
Homo sapiens.
WO2003057926-A1.
17-JUL-2003.
08-JAN-2003; 2003WO-US000657.
08-JAN-2002; 2002US-0345637P.

ALIGNMENTS

RESULT 1
ID ADK11641 standard; DNA; 246 BP.
XX
AC ADK11641;
XX
DT 06-MAY-2004 (first entry)
XX
DE DE
XX
KW Breast cancer differentially expressed gene product #47.
XX
KW ds: cytostatic; gene therapy; DKFZp5661133 activity inhibitor;
XX
OS breast cancer; differential expression.
XX
PN Homo sapiens.
XX
PN WO2003057926-A1.
XX
PD 17-JUL-2003.
XX
PF 08-JAN-2003; 2003WO-US000657.
XX
PR 08-JAN-2002; 2002US-0345637P.

ABS34463 Human liv
ABS09283 Human gen
ABQ20363 Oligonuc1
ABQ20362 Oligonuc1
ABQ57307 Human col
ACN56475 Cotton gy
ADJ40364 Plant cdn
ABV59145 Human pro
ADS48153 Bacterial
ABQ43783 Oligonuc1
ABQ43782 Oligonuc1
ABV12580 Human pro
ADK16832 Nanoarcha
ADC77269 DNA homol
ADK59933 Plant DNA
ADK57152 Plant DNA
ADK57326 Plant DNA
ABT09278 Phase-1 R
ADG31012 Liver tox
ADG45647 Liver inf
ADH22935 Partial D
ADR02380 A. gossyp
ADL63772 Human ova
ABV25474 Human pro
ADR02073 A. gossyp
AA195122 Human neu
AA195567 Human neu
ABQ26378 Oligonuc1
ABQ26379 Oligonuc1
ACH97769 Klebsiell
AAH04768 Human cdn
ACD26286 DNA encod
ABZ15457 Arabidops
ADQ22980 Human sof
ADK58242 Plant DNA
ADK54149 Plant DNA
ABQ66969 Human clo
ADQ25555 Human sof
ADA30629 DNA encod
ADS61127 Bacterial
AAC39924 Arabidops
ACA26583 Prokaryot
AAZ97180 Human pro
ABL14243 Drosophil
ADT05028 Haemophil
ADS49042 Bacterial
ACA48331 Prokaryot

XX (CHIR ) CHIRON CORP.
XX Hansen R;
XX WPI; 2003-577534/54.
XX Inhibiting a cancerous phenotype of a cell, useful for treating breast
cancer comprises contacting a cancerous mammalian cell with an agent for
inhibition of DKFZp5661133 activity.
XX Claim 30; SEQ ID NO 47; 257pp; English.
XX The invention relates to a method of inhibiting a cancerous phenotype of
a cell comprises contacting a cancerous mammalian cell with an agent for
inhibition of DKFZp5661133 activity. The methods are useful for treating
cancer, e.g. breast cancer. This sequence represents a gene product which
is differentially expressed in breast cancer cells. The sequence can be
used in the method of the invention.
XX Sequence 246 BP; 77 A; 49 C; 59 G; 61 T; 0 U; 0 Other;
Alignment Scores: 1.41 Length: 246
Pred. No.: 45.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 10
US-10-774-176-14 (1-9) x ADK11641 (1-246)
QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
|||||
DB 30 GTTTGTATTTCACCGCAAGGGGATA 56
|||||
RESULT 2
ADU11677
ID ADU11677 standard; DNA; 475 BP.
XX
AC ADU11677;
XX
DT 27-JAN-2005 (first entry)
XX
DE Solid tumour prognosis gene seqid 2116.
XX cytostatic; gene therapy; expression profile; solid tumour;
peripheral blood mononuclear cell; PMBC; prognosis; ds.
XX Unidentified.
XX WO2004097052-A2.
XX 11-NOV-2004.
XX 29-APR-2004; 2004WO-US013587.
XX 29-APR-2003; 2003US-0466067P.
XX 23-JAN-2004; 2004US-0538246P.
XX (AMHP ) WYETH.
XX (STRA/) STRAHS A.
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
Immerman F, Dorrner AJ;
XX WPI; 2004-804779/79.
XX A method, useful for prognosing and treating solid tumor, comprises
comparing an expression profile of a gene expressed in peripheral blood
mononuclear cells to a reference expression profile of a gene.
XX Disclosure; Page; 111pp; English.

```

XX The invention describes a method comprising comparing an expression
 CC profile of at least one gene in a peripheral blood sample of a patient to
 CC at least one reference expression profile of the at least one gene, where
 CC the patient has a solid tumour, and each of the gene is differentially
 CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
 CC of patients as compared to PBMCs of a second class of patients, where
 CC both the first and second classes of patients have the solid tumour, and
 CC each of the first and second classes is a subcluster formed by an
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of
 CC a population of patients who have the solid tumour, and where the
 CC majority of the first class of patients has a first clinical outcome, and
 CC the majority of the second class of patients has a second clinical
 CC outcome. Also described are: a system comprising (i) a memory or a
 CC storage medium including data that represent an expression profile of at
 CC least one gene in a peripheral blood sample of a patient who has a solid
 CC tumour, (ii) at least another storage medium including data that
 CC represent at least one reference expression profile of the gene, (iii) a
 CC program capable of comparing the expression profile to the reference
 CC expression profile, and (iv) a processor capable of executing the
 CC program, where expression levels of the gene in peripheral blood
 CC mononuclear cells of patients who have the solid tumour correlate with
 CC clinical outcomes of the patients; and a nucleic acid or protein array
 CC comprising concentrated probes for solid tumour prognosis genes, where
 CC each of the solid tumour prognosis genes is differentially expressed in
 CC PBMCs of a first class of patients as compared to PBMCs of a second class
 CC of patients, where both the first and second classes of patients have a
 CC solid tumour, and where the first class of patients has a first clinical
 CC outcome, and the second class of patients has a second clinical outcome.
 CC The method, system, and array are useful for prognosing and treating
 CC solid tumours. This sequence represents a solid tumour prognosis gene of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.97 Length: 475
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-14 (1-9) x ADU11677 (1-475)

Qy 1 ValLeuTyrLeuAsnArgGlycIyle 9
 Db 396 GTTTTGTATTGACCGCAAGGGGATA 422

RESULT 3
 AAA27060
 ID AAA27060 standard; DNA; 901 BP.

XX AC AAA27060;

XX DT 22-AUG-2000 (first entry)

XX DE Canine 5T4 tumour-associated antigen gene.

XX KW Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
 KW ds.

XX OS Canis sp.

XX FH Key Location/Qualifiers

FT CDS 1..858

FT FT /tag= a

FT FT /product= "5T4 antigen"

FT FT misc_feature 61..74

FT FT /tag= b

FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 135..146
 FT /tag= c
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 207..216
 FT /tag= d
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 277..290
 FT /tag= e
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 351..361
 FT /tag= f
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 422..436
 FT /tag= g
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 497..511
 FT /tag= h
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 572..583
 FT /tag= i
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 644..653
 FT /tag= j
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 714..723
 FT /tag= k
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 784..801
 FT /tag= l
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT WO200029428-A2.
 XX PD 25-MAY-2000.
 XX PF 18-NOV-1999; 99WO-GB003859.
 XX PR 18-NOV-1998; 98GB-00025303.
 XX PR 27-JAN-1999; 99GB-00001739.
 XX PR 30-JUL-1999; 99GB-00017995.
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PI Carroll MW, Myers KA;
 XX WPI; 2000-387735/33.
 XX P-PSDB; AAY94351.
 XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.

XX Disclosure; Page 78-79; 79pp; English.

XX The present sequence encodes the canine 5T4 tumour-associated antigen

CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in

CC carcinomas but has a highly restricted expression pattern in normal adult

CC tissues. It appears to be strongly correlated to metastasis in colorectal

CC and gastric cancer. 5T4 antigen may therefore be useful in tumour

CC diagnosis, targeting and immunotherapy. Mice in which tumours had been

CC induced were inoculated with a virus expression vector containing the

CC human or murine 5T4 gene sequence. The 5T4 antigen was shown to be

CC effective at eliciting an immunotherapeutic anti-tumour response. Both

CC the nucleic acid encoding the antigen and the antigen itself can be used

CC to elicit an immune response, preferably CTL or an antibody response in a

CC subject

XX Sequence 901 BP; 178 A; 246 C; 212 G; 153 T; 0 U; 112 Other;

Alignment Scores:

Pred. No.:	6.13	Length:	901
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

US-10-774-176-14 (1-9) x AAA27060 (1-901)

QY 1 ValLeuTYrLeuAenATcLYsGlyIle 9

Db 684 GTTTTGATTGTAACCCGACGGGATA 710

RESULT 4

ABT07721

ID ABT07721 standard; DNA; 927 BP.

XX AC ABT07721;

XX 14-NOV-2002 (first entry)

XX Breast cancer-associated gene sequence 29.

XX Gene; ds; breast cancer; breast cancer-associated gene sequence;

KW drug development; pharmacogenetics; biosensor development.

XX Unidentified.

XX WO200259377-A2.

XX 01-AUG-2002.

XX 24-JAN-2002; 2002WO-US002242.

XX 24-JAN-2001; 2001US-0263965P.

XX 02-FEB-2001; 2001US-0265928P.

XX 09-APR-2001; 2001US-00829472.

XX 09-APR-2001; 2001US-0282698P.

XX 04-MAY-2001; 2001US-0288590P.

XX 29-MAY-2001; 2001US-0294443P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC, Afar D;

XX WPI; 2002-583738/62.

XX N-PSDB; ABJ05564.

XX Detecting a breast cancer-associated transcript in a patient's cell,

PT useful for diagnosing breast cancer, comprises contacting a biological

PT sample with a polynucleotide that selectively hybridizes with breast

XX cancer nucleic acids.

XX Claim 9; Page 372; 414pp; English.

XX The invention comprises a method of detecting a breast cancer-associated

CC transcript in a cell from a patient. The method of the invention involves

CC contacting a biological sample from the patient with a nucleotide that

CC hybridises to one of the 69 breast cancer-associated gene sequences shown

CC in the specification. The method of the invention is useful in the

CC diagnosis or prognosis of breast cancer, and for detecting genes that are

CC up or down-regulated in breast cancer cells. Genes identified by the

CC method of the invention can be used in diagnostic purposes and also as

CC targets for screening for therapeutic compounds that modulate breast

CC cancer (e.g. hormones or antibodies). Identification of genes that are

CC over or under expressed in breast cancer can additionally provide high-

CC resolution, high-sensitivity datasets which can be used in the areas of

CC diagnostics, therapeutics, drug development, pharmacogenetics, protein

CC structure and biosensor development. DNA sequences ABT07693 - ABT07761

CC represent the 69 breast cancer-associated gene sequences of the invention

XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.33	Length:	927
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-14 (1-9) x ABT07721 (1-927)

QY 1 ValLeuTYrLeuAenATcLYsGlyIle 9

Db 775 GTTTTGATTGTAACCCGACGGGATA 801

RESULT 5

ABX76333

ID ABX76333 standard; DNA; 927 BP.

XX AC ABX76333;

XX 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #197.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;

KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

XX 10-MAY-2001; 2001US-0290492P.

XX 09-NOV-2001; 2001US-0339245P.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0334370P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX P-PSDB; ABU56604.

XX Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased
 XX expression in lung cancer.
 PS Claim 22; Page 336; 453pp; English.
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 6.33 Length: 927
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x ABX76333 (1-927)

QY 1 ValLeuTyrlEuAsnArgLysGlyIle 9
 |||||
 DB 775 GTTTGTATTGACCGCAAGGGGATA 801

RESULT 6
 ADB80503

ID ADB80503 standard; DNA; 927 BP.

XX ADB80503;

XX 04-DEC-2003 (first entry)

XX Ovarian cancer-associated transcript #34.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; ds; gene.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1..927
 FT /*tag= a

XX WO2002102235-A2.

PN 27-DEC-2002.

PD 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

XX 12-APR-2002; 2002US-0372246P.

PA (ROSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

DR P-PSDB; ADB80504.

XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

XX Claim 10; Page 297; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 6.33 Length: 927
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-14 (1-9) x ADB80503 (1-927)

QY 1 ValLeuTyrlEuAsnArgLysGlyIle 9
 |||||
 DB 775 GTTTGTATTGACCGCAAGGGGATA 801

RESULT 7

ADN38723

ID ADN38723 standard; cDNA; 927 BP.

XX ADN38723;

XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

XX WO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

PR 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (BOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KB, Ziolknik A;
 XX
 XX WPI; 2003-468649/44.
 DR P-PSDB; ADN38724.
 DR
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 XX Claim 8; SEQ ID NO 41; 1385pp; English.
 PS
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

 Alignment Scores:
 Pred. No.: 6.33 Length: 927
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

 US-10-774-176-14 (1-9) x ADN38723 (1-927)

 QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 DB 775 GTTTTGATTGACCGCAGGGGATA 801

 RESULT 8
 AAD56198
 ID AAD56198 standard; DNA; 973 BP.
 XX
 AC AAD56198;
 XX
 DT 07-AUG-2003 (first entry)
 XX
 DB Human LRRCAPS related DNA #5.

 XX Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 FN WO2003035831-A2.
 XX
 PD 01-MAY-2003.
 XX
 PF 21-OCT-2002; 2002WO-US033540.
 XX
 PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX
 DR WPI; 2003-421410/39.
 XX
 PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 PS Example 5; Page 74-75; 99pp; English.
 XX
 CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX
 SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

 Alignment Scores:
 Pred. No.: 6.69 Length: 973
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

 US-10-774-176-14 (1-9) x AAD56198 (1-973)

 QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 DB 790 GTTTTGATTGACCGCAGGGGATA 816

 RESULT 9
 ABV99349
 ID ABV99349 standard; DNA; 1156 BP.
 XX
 AC ABV99349;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Human NOV8a coding sequence.
 XX
 KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
 KW neutropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;

KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
XX cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
OS Homo sapiens.
XX W0200272771-A2.
XX 19-SEP-2002.
XX 08-MAR-2002; 2002WO-US007288.
XX 08-MAR-2001; 2001US-0274101P.
XX 08-MAR-2001; 2001US-0274194P.
XX 08-MAR-2001; 2001US-0274281P.
XX 08-MAR-2001; 2001US-0274322P.
XX 09-MAR-2001; 2001US-0274849P.
XX 12-MAR-2001; 2001US-0275235P.
XX 13-MAR-2001; 2001US-0275578P.
XX 13-MAR-2001; 2001US-0275579P.
XX 13-MAR-2001; 2001US-0275601P.
XX 14-MAR-2001; 2001US-0276000P.
XX 16-MAR-2001; 2001US-0276776P.
XX 19-MAR-2001; 2001US-0276994P.
XX 20-MAR-2001; 2001US-0277239P.
XX 20-MAR-2001; 2001US-0277321P.
XX 20-MAR-2001; 2001US-0277327P.
XX 20-MAR-2001; 2001US-0277389P.
XX 21-MAR-2001; 2001US-0277791P.
XX 22-MAR-2001; 2001US-0277833P.
XX 23-MAR-2001; 2001US-0278152P.
XX 26-MAR-2001; 2001US-0278894P.
XX 27-MAR-2001; 2001US-0278999P.
XX 27-MAR-2001; 2001US-0279036P.
XX 28-MAR-2001; 2001US-0279344P.
XX 30-MAR-2001; 2001US-0279995P.
XX 30-MAR-2001; 2001US-0280233P.
XX 02-APR-2001; 2001US-0280802P.
XX 02-APR-2001; 2001US-0280822P.
XX 02-APR-2001; 2001US-0280900P.
XX 04-APR-2001; 2001US-0281194P.
XX 13-APR-2001; 2001US-0283675P.
XX 30-APR-2001; 2001US-0287424P.
XX 02-MAY-2001; 2001US-0288066P.
XX 03-MAY-2001; 2001US-0288342P.
XX 15-MAY-2001; 2001US-0288528P.
XX 15-MAY-2001; 2001US-0291190P.
XX 16-MAY-2001; 2001US-0291099P.
XX 16-MAY-2001; 2001US-0291240P.
XX 30-MAY-2001; 2001US-029485P.
XX 31-MAY-2001; 2001US-0294889P.
XX 31-MAY-2001; 2001US-0294899P.
XX 18-JUN-2001; 2001US-0299027P.
XX 19-JUN-2001; 2001US-0299303P.
XX 10-JUL-2001; 2001US-0304354P.
XX 31-JUL-2001; 2001US-0309198P.
XX 16-AUG-2001; 2001US-0312903P.
XX 10-SEP-2001; 2001US-0318462P.
XX 12-SEP-2001; 2001US-0318770P.
XX 27-SEP-2001; 2001US-0325430P.
XX 18-OCT-2001; 2001US-0325681P.
XX 31-OCT-2001; 2001US-0330380P.
XX 14-NOV-2001; 2001US-0335301P.
XX 14-NOV-2001; 2001US-0332271P.
XX 14-NOV-2001; 2001US-0332271P.
XX 14-NOV-2001; 2001US-0332272P.
XX 14-NOV-2001; 2001US-0333184P.
XX 21-NOV-2001; 2001US-0333272P.
XX 03-DEC-2001; 2001US-0332094P.
XX 03-DEC-2001; 2001US-0337426P.
XX 04-DEC-2001; 2001US-0338092P.
XX 04-DEC-2001; 2001US-0337185P.

PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX (CURA-) CURAGEN CORP.
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
PI Boldog PL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Pochart PP;
PI Zhong M;
XX WPI; 2002-732824/79.
DR P-PSDB; ABP70071.
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
PT disorders, and asthma.
XX Claim 16; Page 114-115; 619pp; English.
XX The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABP70043-ABP70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, haematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 8.13 Length: 1156
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x ABV99349 (1-1156)
QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
|||||
Db 1006 GTTTGTATTTCACCGCAGGGGATA 1032
RESULT 10
ABK87175
ID ABK87175 standard; cDNA; 1260 BP.
XX ABK87175;
XX 07-OCT-2002 (first entry)
XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 574.
XX Feline; cat; oncofoetal leucine-rich glycoprotein; 574; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX

OS Felis sp.
 XX Key Location/Qualifiers
 XX CDS 1..1260
 XX FT /*tag= a
 XX FT /product= "5T4 protein"
 XX PN WO200238612-A2.
 XX 16-MAY-2002.
 XX 13-NOV-2001; 2001WO-GB005004.
 XX 13-NOV-2000; 2000WO-GB004317.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Myers K, Drury N, Carroll M;
 XX WPI; 2002-557449/59.
 XX P-PSDB; AAU98694.
 XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 XX polypeptide, useful in preparation of vaccine for treating and/or
 XX preventing cancer in a subject, preferably a dog or cat.
 XX Claim 4; Page 68; 68pp; English.
 XX The present invention relates to the isolation of canine and feline
 XX oncofoetal leucine-rich glycoproteins known as 5T4, and the
 XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 XX a significant proportion of tumours. The sequences of the invention are
 XX useful in a pharmaceutical composition for the prevention and/or
 XX treatment of tumours or other diseases associated with cell
 XX proliferation, infections, and inflammatory conditions in animals,
 XX preferably dogs or cats. The compositions may also be used for cancer
 XX immunotherapy in these animals. The sequences of the invention may also
 XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
 XX measurement and localisation of 5T4 in extracts of plasma, urine,
 XX tissues, and in cell culture media. Antibodies specific for the 5T4
 XX protein are useful for isolating foetal cells from maternal blood. The
 XX isolation process may form part of a diagnostic method e.g. the foetal
 XX cells may then be subject to biochemical or genetic sampling used for
 XX testing foetal abnormalities, or to determine the sex of the foetus(es).
 XX The present sequence encodes feline 5T4 protein
 XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 8.96 Length: 1260
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-14 (1-9) x ABR87175 (1-1260)
 Qy 1 ValLeuTYrLeuAenArgLYsGLyIle 9
 Db 1114 GTTTGTACTTGAACCGCAGGGGATA 1140
 RESULT 11
 ADB97513
 ID ADB97513 standard; DNA; 1260 BP.
 XX AC ADB97513;
 XX 04-DEC-2003 (first entry)
 XX Feline 5T4 antigen DNA.
 XX Major Histocompatibility Complex class I peptide epitope; MHC;
 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 cyostatic; cancer; feline; gene; ds.
 Unidentified.
 Key Location/Qualifiers
 CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"
 WO2003068816-A1.
 21-AUG-2003.
 13-FEB-2003; 2003WO-GB000670.
 13-FEB-2002; 2002GB-00003419.
 (OXFO-) OXFORD BIOMEDICA UK LTD.
 Carroll M, Kingsman S, Redchenko I;
 WPI; 2003-637141/60.
 P-PSDB; ADB97520.
 New major histocompatibility complex class I peptide epitopes from human
 5T4 tumor-associated antigen, useful for preventing and/or treating a
 disease, particularly cancer.
 Disclosure; Page 67; 73pp; English.
 The invention relates to a novel Major Histocompatibility Complex (MHC)
 class I peptide epitope of the 5T4 antigen. The invention further
 provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 epitope; a vector system capable of delivering the 5T4 epitope nucleic
 acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 comprising the above; a method for treating and/or preventing a disease
 in a subject by administering the vaccine; an agent capable of binding
 specifically to the 5T4 epitope and/its encoding nucleic acid; a method
 comprising detecting the presence of the 5T4 epitope or its encoding
 nucleic acid in a subject; and a T cell line or clone capable of
 specifically recognising the 5T4 epitope in conjunction with an MHC class
 I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 comprising the 5T4 epitope or its encoding nucleic acid and the vector
 system or cell is useful in the prevention and/or treatment of a disease,
 particularly cancer. The detection method is useful for diagnosing or
 monitoring the progression of a cancerous disease, and for detecting the
 presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 is useful in the manufacture of a medicament for treating and/or
 preventing a disease. This polynucleotide sequence represents the feline
 5T4 antigen coding DNA of the invention.
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 8.96 Length: 1260
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-14 (1-9) x ADB97513 (1-1260)
 Qy 1 ValLeuTYrLeuAenArgLYsGLyIle 9
 Db 1114 GTTTGTACTTGAACCGCAGGGGATA 1140
 RESULT 12
 ADB97452
 ID ADB97452 standard; DNA; 1260 BP.

```
XX AC ADB97452;
XX DT 04-DEC-2003 (first entry)
XX DE DNA encoding feline 5T4 protein.
XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
XX KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
XX OS Unidentified.
XX FH Key Location/Qualifiers
XX FT CDS 1..1260
XX FT /*tag= a
XX FT /product= "Feline 5T4 antigen protein"
XX PN WO2003068815-A2.
XX PD 21-AUG-2003.
XX PF 13-FEB-2003; 2003WO-GB000618.
XX PR 13-FEB-2002; 2002GB-00003420.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll M, Harrop R, Kingsman S;
XX PS WPI; 2003-663795/62.
XX DR P-PSDB; ADB97455.
XX KW New Major Histocompatibility Complex class II peptide epitope of 5T4,
XX PT useful for manufacturing a medicament for diagnosing, preventing and/or
XX PT treating a disease, e.g. cancer.
XX PS Disclosure; Page 49; 63pp; English.
XX CC The invention relates to a Major Histocompatibility Complex (MHC) class
XX CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
XX CC clone has a cytostatic activity, as it is useful in manufacturing a
XX CC medicament for preventing and/or treating a disease, particularly cancer.
XX CC The methods are useful for detecting T-cells capable of specifically
XX CC recognising a peptide epitope in conjunction with an MHC molecule, for
XX CC diagnosing or monitoring the progression of a cancerous disease, or for
XX CC detecting the presence of a peptide or nucleic acid using an agent. The
XX CC MHC class II peptide epitope of the invention can be used in gene therapy
XX CC or as part of a vaccine. This polynucleotide sequence represents the DNA
XX CC coding for the feline 5T4 protein.
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. NO.: 8.96 Length: 1260
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-14 (1-9) x ADB97452 (1-1260)

Qy 1 ValLeuTyrlLeuAsnArgLysGlyIle 9
Db 1114 GTTTGTACTTGAACCGCAAGGGGATA 1140

RESULT 13
AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX AC AAA27058;
XX DT 22-AUG-2000 (first entry)
```

```
XX DE Human 5T4 tumour-associated antigen gene.
XX KW Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX KW ds.
XX OS Homo sapiens.
XX PN WO200029428-A2.
XX PD 25-MAY-2000.
XX PF 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MW, Myers KA;
XX PS WPI; 2000-387735/33.
XX KW Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX PT response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the human 5T4 tumour-associated antigen
XX CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The 5T4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject
XX SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 8.99 Length: 1263
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-14 (1-9) x AAA27058 (1-1263)

Qy 1 ValLeuTyrlLeuAsnArgLysGlyIle 9
Db 1117 GTTTGTATTTCACCGCAAGGGGATA 1143

RESULT 14
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX AC AAF89736;
XX DT 23-JUL-2001 (first entry)
XX DE Nucleotide sequence of canine 5T4 protein.
XX KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
XX KW hypersensitivity; autoimmune disease; central nervous system disease;
XX KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
XX KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
XX KW Helicobacter-related disease; immune disorder; ss.
```

XX Canis sp.
 XX Key Location/Qualifiers
 XX CDS 1..1263
 XX /*tag= a
 XX /product= "5T4"
 XX WO200136486-A2.
 XX 25-MAY-2001.
 XX 13-NOV-2000; 2000WO-GB004317.
 XX 18-NOV-1999; 99WO-GB003859.
 XX 15-FEB-2000; 2000GB-00003527.
 XX 02-MAR-2000; 2000GB-00005071.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
 XX Myers KA;
 XX WPI; 2001-343805/36.
 XX P-PSDB; AAB83839.
 XX Use of single chain antibody capable of recognizing a disease associated
 XX molecule for manufacturing a medicament for preventing and/or treating a
 XX disease condition associated with disease associated molecule.
 XX Disclosure; Fig 26; 118pp; English.
 XX The specification describes the use of a single chain antibody (ScFv),
 XX which is capable of recognizing a disease associated molecule in the
 XX manufacture of a medicament for the prevention and treatment of a disease
 XX condition. The ScFv antibody is useful in the manufacture of a
 XX medicament, for affecting a disease in vivo, for preparing a
 XX pharmaceutical composition, for in vivo imaging and/or for adjuvant
 XX treatment of a disease. The ScFv antibody is also useful for treating
 XX inflammatory diseases including arthritis, hypersensitivity, autoimmune
 XX diseases, cancers, central nervous system disorders including Parkinson's
 XX disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
 XX diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
 XX related diseases, and other immune disorders. The present sequence
 XX encodes a 5T4 protein, which is used to produce ScFv of the invention
 XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8.99 Length: 1263
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-14 (1-9) x AAF89736 (1-1263)
 Qy 1 ValLeuTyrLeuAenArgLysGlyIle 9
 Db 1117 GTTTTGTATTGTAACCGCAAGGGGATA 1143
 RESULT 15
 ABK87174
 ID ABK87174 standard; cDNA; 1263 BP.
 XX
 XX ABK87174;
 AC
 XX 07-OCT-2002 (first entry)
 DT
 XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
 DE
 XX Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW

KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX Canis sp.
 XX Key Location/Qualifiers
 XX CDS 1..1263
 XX /*tag= a
 XX /product= "5T4 protein"
 XX WO200238612-A2.
 XX 16-MAY-2002.
 XX 13-NOV-2001; 2001WO-GB005004.
 XX 13-NOV-2000; 2000WO-GB004317.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Myers K, Drury N, Carroll M;
 XX WPI; 2002-557449/59.
 XX P-PSDB; AAU98693.
 XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 XX polypeptide, useful in preparation of vaccine for treating and/or
 XX preventing cancer in a subject, preferably a dog or cat.
 XX Claim 1; Page 67; 68pp; English.
 XX The present invention relates to the isolation of canine and feline
 XX oncofoetal leucine-rich glycoproteins known as 5T4, and the
 XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 XX a significant proportion of tumours. The sequences of the invention are
 XX useful in a pharmaceutical composition for the prevention and/or
 XX treatment of tumours or other diseases associated with cell
 XX proliferation, infections, and inflammatory conditions in animals,
 XX preferably dogs or cats. The compositions may also be used for cancer
 XX immunotherapy in these animals. The sequences of the invention may also
 XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
 XX measurement and localisation of 5T4 in extracts of plasma, urine,
 XX tissues, and in cell culture media. Antibodies specific for the 5T4
 XX protein are useful for isolating foetal cells from maternal blood. The
 XX isolation process may form part of a diagnostic method e.g. the foetal
 XX cells may then be subject to biochemical or genetic sampling used for
 XX testing foetal abnormalities, or to determine the sex of the foetus(es).
 XX The present sequence encodes canine 5T4 protein
 XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8.99 Length: 1263
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-14 (1-9) x ABK87174 (1-1263)
 Qy 1 ValLeuTyrLeuAenArgLysGlyIle 9
 Db 1117 GTTTTGTATTGTAACCGCAAGGGGATA 1143
 RESULT 16
 AAA27059
 ID AAA27059 standard; DNA; 1261 BP.
 XX
 XX AAA27059;
 AC
 XX 22-AUG-2000 (first entry)
 DT

XX Mouse 5T4 tumour-associated antigen gene.
DE
XX Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
XX Mus musculus.
OS
XX WO200029428-A2.
PN
XX 25-MAY-2000.
PD
XX 18-NOV-1999; 99WO-GB003859.
PF
XX 18-NOV-1998; 98GB-00025303.
PR
XX 27-JAN-1999; 99GB-00001739.
PR
XX 30-JUL-1999; 99GB-00017995.
PR
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX Carroll MW, Myers KA;
PI
XX WPI; 2000-387735/33.
XX
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
PT response useful in vaccinating against and in treating tumors.
PT
XX
XX Example 2; Page 78; 79pp; English.
PS
XX The present sequence encodes the mouse 5T4 tumour-associated antigen
CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
CC carcinomas but has a highly restricted expression pattern in normal adult
CC tissues. It appears to be strongly correlated to metastasis in colorectal
CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
CC induced were inoculated with a virus expression vector containing the
CC present sequence. The 5T4 antigen was shown to be effective at eliciting
CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
CC the antigen and the antigen itself can be used to elicit an immune
CC response, preferably CTL or an antibody response in a subject. The
CC present sequence appears in GenBank at accession number AJ012160
XX
SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.13 Length: 1281
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-14 (1-9) x AAA27059 (1-1281)

QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
Db 1135 GTTTTGTATTGTAACCGTAAGGCATA 1161

RESULT 17
AAD56199
ID AAD56199 standard; DNA; 1331 BP.
XX
AC AAD56199;
XX
XX 07-AUG-2003 (first entry)
DT
XX Human LRRCAPS related DNA #6.
DE
XX Human; p53 pathway; leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
KW
XX Homo sapiens.
OS

XX WO2003035831-A2.
PN
XX 01-MAY-2003.
PD
XX 21-OCT-2002; 2002WO-US033540.
PF
XX 22-OCT-2001; 2001US-0338733P.
PR
XX 15-FEB-2002; 2002US-0357600P.
PR
XX 01-MAR-2002; 2002US-0361196P.
PR
XX (EXEL-) EXELIXIS INC.
PA
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
PI
XX WPI; 2003-421410/39.
DR
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
PT
XX Disclosure; Page 75-76; 99pp; English.
XX
XX The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.54 Length: 1331
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x AAD56199 (1-1331)

QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
Db 1147 GTTTTGTATTGTAACCGCAAGGCATA 1173

RESULT 18
AAD56299
ID ADJ56299 standard; cDNA; 2020 BP.
XX
AC ADJ56299;
XX
XX 06-MAY-2004 (first entry)
DT
XX Human cDNA differentially expressed in MYCN activated cells SeqID 105.
DE
XX human; differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX
XX Homo sapiens.
OS
XX US2003119009-A1.
PN
XX 26-JUN-2003.
PD
XX 25-FEB-2002; 2002US-00084817.
XX
XX

PR 23-FEB-2001; 2001US-0270784P.
XX
PA (STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLON/) PLON S E.
PA (SHOH/) SHOHET J M.
XX
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX
DR WPI; 2003-635698/60.
XX
PT New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
PS Claim 1; SEQ ID NO 105; 27pp; English.
XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX
SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 15.3 Length: 2020
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-14 (1-9) x ADJ56299 (1-2020)
Qy 1 ValLeuTyrlLeuAenArgLysGlyIle 9
Db 1187 GTTTTGTATTGTAACCGCAGGGGATA 1213
RESULT 19
ACCS1052
ID ACCS1052 standard; cDNA; 2053 BP.
XX
AC ACCS1052;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated cDNA sequence SEQ ID NO:192.
KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
OS
XX WO2003003906-A2.
XX
XX 16-JAN-2003.
XX
XX 03-JUL-2002; 2002WO-US021338.
XX
XX 03-JUL-2001; 2001US-0302814P.

PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Aziz N;
XX
DR WPI; 2003-201532/19.
DR P-PSDB; ABR48236.
XX
PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
PS Claim 6; Page 296; 307pp; English.
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridises to a sequence that is 80 % identical to a
CC table of sequences (see ACCS0951 to ACCS1059). ACCS0951 to ACCS1059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 15.6 Length: 2053
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x ACCS1052 (1-2053)
Qy 1 ValLeuTyrlLeuAenArgLysGlyIle 9
Db 1201 GTTTTGTATTGTAACCGCAGGGGATA 1227
RESULT 20
ABX76332
ID ABX76332 standard; DNA; 2053 BP.
XX
AC ABX76332;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polynucleotide #196.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
XX WO200286443-A2.
XX
XX 31-OCT-2002.
XX
XX 18-APR-2002; 2002WO-US012476.

XX PR 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX FA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PI Aziz N, Murray R;
 XX DR WPI; 2003-093161/08.
 XX DR P-PSDB; ABUS6603.
 XX PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 XX expression in lung cancer.
 XX PS Claim 22; Page 335; 453pp; English.
 XX CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 XX invention
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-14 (1-9) x ABX76332 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTTGATTTCGACCGCAAGGGGATA 1227
 RESULT 21
 AAD56197
 ID AAD56197 standard; DNA; 2053 BP.
 XX AC AAD56197;
 XX DT 07-AUG-2003 (first entry)
 XX DE Human LRRCAPS DNA #11.
 XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 XX KW LRRCAPS; cancer; gene therapy; ds.
 XX OS Homo sapiens.

PN WO2003035831-A2.
 XX PD 01-MAY-2003.
 XX PF 21-OCT-2002; 2002WO-US033540.
 XX PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX (EXEL-) EXELIXIS INC.
 XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX WPI; 2003-421410/39.
 XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX Example 5; Page 73-74; 99pp; English.
 XX CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-14 (1-9) x AAD56197 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTTGATTTCGACCGCAAGGGGATA 1227
 RESULT 22
 AAD56200
 ID AAD56200 standard; DNA; 2053 BP.
 XX AC AAD56200;
 XX DT 07-AUG-2003 (first entry)
 XX DE Human LRRCAPS DNA #12.
 XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 XX KW LRRCAPS; cancer; gene therapy; ds.
 XX OS Homo sapiens.
 XX PN WO2003035831-A2.
 XX PD 01-MAY-2003.
 XX PF 21-OCT-2002; 2002WO-US033540.
 XX PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.

PR 01-MAR-2002; 2002US-0361196P.
 XX (EXEL-) EXELIXIS INC.
 XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 PI WPI; 2003-421410/39.
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX Disclosure; Page 76-77; 99pp; English.
 XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-14 (1-9) x AAD56200 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTTGTATTGTAACCGCAGGGGATA 1227
 RESULT 23
 ADN38721
 ID ADN38721 standard; cDNA; 2053 BP.
 XX
 AC ADN38721;
 DT 17-JUN-2004 (first entry)
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.
 XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulneryary; gene therapy; vaccine; gene; ss.
 XX Homo sapiens.
 XX WO2003042661-A2.
 PN 22-MAY-2003.
 PD 13-NOV-2002; 2002WO-US036810.
 XX 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0352505P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396899P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 PI WPI; 2003-468649/44.
 DR P-PSDB; ADN38722.
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX Claim 8; SEQ ID NO 39; 1385pp; English.
 XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0
 US-10-774-176-14 (1-9) x ADN38721 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTTGTATTGTAACCGCAGGGGATA 1227
 RESULT 24
 ADL06473
 ID ADL06473 standard; cDNA; 2053 BP.
 XX
 AC ADL06473;
 XX 20-MAY-2004 (first entry)
 DT Human tumour-associated antigenic target (TAT) cDNA sequence #53.
 DE Human; tumour-associated antigenic target; TAT; cell death; tumour;
 KW

KW cancer; cytostatic; gene; ss.
 XX OS Homo sapiens.
 XX WO2004016225-A2.
 XX PN 26-FEB-2004.
 XX PD
 XX PA (GETH) GENENTECH INC.
 XX PF 19-AUG-2003; 2003WO-US025892.
 XX PR 19-AUG-2002; 2002US-0404809P.
 XX PR 21-AUG-2002; 2002US-0405645P.
 XX PR 23-SEP-2002; 2002US-0413192P.
 XX PR 15-OCT-2002; 2002US-0419008P.
 XX PR 15-NOV-2002; 2002US-0426847P.
 XX PR 02-JUL-2003; 2003US-0484959P.
 XX PA (GETH) GENENTECH INC.
 XX PI Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
 PI Spencer SD, Wu TD, Zhang Z;
 XX WPI: 2004-257144/24.
 XX P-PSDB; ADL06552.
 XX New antibody that binds to a tumor-associated antigenic target (TAT)
 PT polypeptide, useful for preparing a composition for diagnosing or
 PT treating cancer.
 XX Claim 1; SEQ ID NO 53; 319pp; English.
 XX The present invention relates to the isolation of human tumour-associated
 CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also
 CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is
 CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a
 CC humanised antibody. It is conjugated to a growth inhibitory agent. It is
 CC produced in bacteria or in CHO cells and induces death of a cell to which
 CC it binds. The antibody is useful for preparing a composition for
 CC diagnosing or treating tumours and cancer. The present sequence
 CC represents a human TAT cDNA sequence of the invention.
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-14 (1-9) x ADL06473 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTGTATTGACCGCAAGGGGATA 1227
 RESULT 25
 ADN03961
 ID ADN03961 standard; cDNA; 2053 BP.
 XX AC ADN03961;
 XX 01-JUL-2004 (first entry)
 XX DE Antipsoriatic cDNA sequence #180.
 XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
 XX OS Homo sapiens.
 XX WO2004028479-A2.
 XX PN
 XX PT

PD 08-APR-2004.
 XX 25-SEP-2003; 2003WO-US030907.
 XX 25-SEP-2002; 2002US-0414006P.
 XX PA (GETH) GENENTECH INC.
 XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 PI Wu TD;
 XX WPI: 2004-305105/28.
 XX P-PSDB; ADN03962.
 XX New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 XX Claim 1; SEQ ID NO 355; 3069pp; English.
 XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polynucleotides of the invention.
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-14 (1-9) x ADN03961 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTGTATTGACCGCAAGGGGATA 1227
 RESULT 26
 ADR25444
 ID ADR25444 standard; DNA; 2053 BP.
 XX AC ADR25444;
 XX 21-OCT-2004 (first entry)
 XX DE Breast cancer prognosis marker #1305.
 XX ds; breast cancer; prognosis; gene expression; diagnosis.
 XX OS Homo sapiens.
 XX WO2004065545-A2.
 XX 05-AUG-2004.
 XX 15-JAN-2004; 2004WO-US001100.
 XX 15-JAN-2003; 2003US-00342887.
 XX (ROSE-) ROSETTA INPHARMATICS LLC.
 XX PA (NECA-) NETHERLANDS CANCER INST.
 XX Van't Veer LJ, He Y;
 XX WPI: 2004-593473/57.
 XX Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of

PT five genes in a cell sample taken from patient, to control levels.
 XX Disclosure; SEQ ID NO 1305; 226pp; English.
 CC The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-14 (1-9) x ADR25444 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTTGTATTGACCCGAGGGGATA 1227
 RESULT 27
 ACN38510
 ID ACN38510 standard; cDNA; 2053 BP.
 XX
 AC ACN38510;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Tumour-associated antigenic target (TAT) cDNA DNA103471, SEQ ID NO:2070.
 XX
 KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO2004030615-A2.
 XX
 PD 15-APR-2004.
 XX
 PF 29-SEP-2003; 2003WO-US028547.
 XX
 PR 02-OCT-2002; 2002US-0414971P.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Wu TD, Zhang Z, Zhou Y;
 PI
 XX
 DR WPI; 2004-347921/32.
 DR
 DR P-PSDB; ABM80804.
 XX
 XX New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.
 XX
 PS Claim 1; SEQ ID NO 2070; 7273pp; English.
 XX

CC The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT nucleic acid of the invention
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.6 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-14 (1-9) x ACN38510 (1-2053)
 Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 Db 1201 GTTTTGTATTGACCCGAGGGGATA 1227
 RESULT 28
 ADV35098
 ID ADV35098 standard; cDNA; 2053 BP.
 XX
 AC ADV35098;
 XX
 DT 10-FEB-2005 (first entry)
 XX
 DE Human cDNA of an exemplary efficacy gene for BAD SeqID174.
 XX
 KW human; ss; multi-parameter high throughput screening; MPHTS;
 KW disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
 KW bipolar affective disorder; BAD; autism; Parkinson's;
 KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.
 XX
 OS Homo sapiens.
 XX
 XX US2003096264-A1.
 XX
 PD 22-MAY-2003.
 XX
 PF 18-JUN-2002; 2002US-00175523.
 XX
 PR 18-JUN-2001; 2001US-0299151P.
 PR 07-SEP-2001; 2001US-0317828P.
 PR 25-SEP-2001; 2001US-0325150P.
 PR 14-NOV-2001; 2001US-0333047P.
 PR 18-JAN-2002; 2002US-0349936P.
 PR 04-MAR-2002; 2002US-0361834P.
 XX
 XX (PSYC-) PSYCHIATRIC GENOMICS INC.
 PA
 XX Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
 PI Palfreyman M, Rajan P;
 XX
 XX WPI; 2004-118903/12.
 DR

XX Identifying a compound that can treat disease or disorders, such as, a
PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
PT determining the expression of one or more efficacy genes in a cell
PT contacted with the test compound.
XX
PS Example 6; SEQ ID NO 174; 39pp; English.
XX
CC This invention relates to a novel screening method identified as a multi-
CC parameter high throughput screening (MPTS) assay. Specifically, it
CC refers to an assay that utilizes the disease signature of a plurality of
CC specific genes associated with a particular disease, and identifies
CC differential expression between those cells taken from individuals
CC affected by that disease and those that are not affected. The present
CC invention then describes the screening of candidate pharmaceutical
CC compounds to identify those that have a potential therapeutic benefit for
CC the treatment of neuropsychiatric and neurodegenerative disorders
CC including schizophrenia, bipolar affective disorder (BAD) and autism, as
CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
CC of this invention exhibit various activities including neuroleptic,
CC nootropic, antimanic and antidepressant. Furthermore, the screening
CC method used in MPTS will be automated, such that a large number of test
CC compounds may be rapidly screened with a minimal amount of labour and
CC effort. This polynucleotide is a human cDNA sequence of a gene that is
CC differentially expressed in the presence of a therapeutic compound and
CC represents an exemplary efficacy gene for bipolar affective disorder,
CC given in an exemplification of the invention.
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 15.6 Length: 2053
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0
US-10-774-176-14 (1-9) x ADV35098 (1-2053)
Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
Db 1201 GTTTTGATTTCGACCGCAAGGGGATA 1227
RESULT 29
AAS87175
ID AAS87175 standard; cDNA; 2338 BP.
AC AAS87175;
DT 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #22979.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSB-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI, 2001-639362/73.

DR P-PSDB; ABC22988.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 22979; 103pp; English.
PS
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 18.1 Length: 2338
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0
US-10-774-176-14 (1-9) x AAS87175 (1-2338)
Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
Db 1458 GTTTTGATTTCGACCGCAAGGGGATA 1484
RESULT 30
AAS94253
ID AAS94253 standard; cDNA; 2359 BP.
XX
XX AAS94253;
XX
XX 06-NOV-2001 (first entry)
DT
XX Human full-length cDNA, SEQ ID NO: 2864.
DE
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX
XX 11-JAN-2000; 2000JP-0018774.
XX
XX 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR P-PSDB; AAM93333.
 XX
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX
 PS Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX
 SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 18.3 Length: 2359
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-14 (1-9) x AAK94253 (1-2359)

QY 1 ValLeuTyxLeuAnArgLysGlyIle 9
 DB 1540 GTTTGTATTGAACCGCAAGGGGATA 1566

RESULT 31
 ADL30831
 ID ADL30831 standard; cDNA; 2359 BP.

XX AC ADL30831;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 XX Full length human cDNA clone SeqID 2864.
 DE
 DE human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.
 XX
 OS Homo sapiens.
 XX
 PN EP1396543-A2.
 XX
 PD 10-MAR-2004.
 XX
 XX 07-JUL-2000; 2003EP-00025638.
 PF
 PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2004-204755/20.
 DR P-PSDB; ADL30832.

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX
 PS Example 1; SEQ ID NO 2864; 1340pp; English.
 XX
 CC This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction.
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.
 XX
 SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 18.3 Length: 2359
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-14 (1-9) x ADL30831 (1-2359)

QY 1 ValLeuTyxLeuAnArgLysGlyIle 9
 DB 1540 GTTTGTATTGAACCGCAAGGGGATA 1566

RESULT 32
 AAK94254

ID AAK94254 standard; cDNA; 2361 BP.

XX AC AAK94254;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 XX Human full-length cDNA, SEQ ID NO: 2866.
 DE
 DE Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 KW
 XX Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 XX 05-SEP-2001.
 PD
 XX 07-JUL-2000; 2000EP-00114089.
 PF
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX
 PA (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR P-PSDB; AAM93334.

XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.

XX Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful

for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO

Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 18.3 Length: 2361
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-14 (1-9) x AAK94254 (1-2361)

Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9
Db 1542 GTTTGTATTGACCGCAGGGGATA 1568

RESULT 33
ADI26162
ID ADI26162 standard; cDNA; 2361 BP.
XX
AC ADI26162;

XX 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #64.

XX ss; gene; human; signal transducer and activator of transcription 6;
KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

XX Homo sapiens.

XX WO2003104277-A2.

XX 18-DEC-2003.

XX 05-JUN-2003; 2003WO-JP007123.

XX 05-JUN-2002; 2002JP-00164257.

XX 06-JUN-2002; 2002US-0385912P.

XX 26-DEC-2002; 2002JP-00377326.

XX 27-DEC-2002; 2002US-0436467P.

XX 15-MAY-2003; 2003JP-00137505.

XX 16-MAY-2003; 2003US-0470836P.

XX (ASAH) ASahi KASEI KK.

XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX WPI; 2004-122214/12.

XX P-PSDB; ADI26163.

XX New signal transducer and activator of transcription 6 activation

PT promoting purified protein, for diagnosing and treating disease

PT associated with activation/inhibition of transcription factor e.g.

PT diabetes and cancer.

XX Claim 4; SEQ ID NO 127; 1368pp; English.

XX The invention relates to a purified protein promoting signal transducer

CC and activator of transcription 6 activation (STAT6). The protein is

CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infections disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficiently promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.

XX Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 18.3 Length: 2361
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-14 (1-9) x ADI26162 (1-2361)

Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9
Db 1542 GTTTGTATTGACCGCAGGGGATA 1568

RESULT 34

ADL30833

ID ADL30833 standard; cDNA; 2361 BP.

XX AC ADL30833;

XX 20-MAY-2004 (first entry)

XX Full length human cDNA clone SeqID 2866.

XX human; medicine; signal transduction; glycoprotein; transcription;

KW oligo-capping method; ss; gene.

XX Homo sapiens.

XX EP1396543-A2.

XX 10-MAR-2004.

XX 07-JUL-2000; 2003EP-00025638.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183865.

XX 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

XX P-PSDB; ADL30834.

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full

PT length human cDNAs.

XX Example 1; SEQ ID NO 2866; 1340pp; English.

XX This invention relates to a novel primers useful for synthesizing full

CC length cDNA molecules that encode human proteins. Specifically, it refers

CC to secretory or membrane proteins that are potential therapeutic agents/

CC target molecules in the field of medicine, and in particular genes

CC encoding proteins that are associated with signal transduction,

CC glycoproteins and transcription. The present invention describes a method

CC for efficiently cloning a full length human cDNA from both the 5' and 3'

CC ends using the oligo-capping method. This polynucleotide sequence is a

CC full length human cDNA clone of the invention.

XX Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	18.3	Length:	2361
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-14 (1-9) x ADL30833 (1-2361)

Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9

Db 1542 GTTTGTATTGACCGAAGGGGATA 1568

RESULT 35

AD126160

ID AD126160 standard; cDNA; 2557 BP.

XX AC AD126160;

XX 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #63.

ss; gene; human; signal transducer and activator of transcription 6;

STAT6; immunogen; STAT6 activation; allergy; inflammation;

autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;

Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;

systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

OS Homo sapiens.

XX WO2003104277-A2.

XX 18-DEC-2003.

XX 05-JUN-2003; 2003WO-JP007123.

XX 05-JUN-2002; 2002JP-00164257.

XX 06-JUN-2002; 2002US-0385912P.

XX 26-DEC-2002; 2002JP-00377326.

XX 27-DEC-2002; 2002US-0436467P.

XX 15-MAY-2003; 2003JP-00137505.

XX 16-MAY-2003; 2003US-0470836P.

XX (ASAH) ASAH KASEI KK.

XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

PI WPI; 2004-122214/12.

XX P-PSDB; AD126161.

XX New signal transducer and activator of transcription 6 activation

PT promoting purified protein, for diagnosing and treating disease

PT associated with activation/inhibition of transcription factor e.g.

PT diabetes and cancer.

XX

PS Claim 4; SEQ ID NO 125; 1368pp; English.

XX The invention relates to a purified protein promoting signal transducer

CC and activator of transcription 6 activation (STAT6). The protein is

CC useful for the producing an antibody, which involves administering the

CC protein or its epitope-bearing fragments to a non-human animal as an

CC antigen. The nucleic acid is useful for diagnosing a disease or

CC susceptibility to a disease related to expression or activity of the

CC protein. A transformant expressing the protein is useful for screening

CC compounds which inhibit or promote STAT6 activation. A transformant

CC expressing the protein is useful for producing a pharmaceutical

CC composition. Compositions, antibodies and antisense molecules are useful

CC for the treating a disease associated with STAT6 activation such as

CC allergic diseases, inflammation, autoimmune diseases, diabetes,

CC hyperlipidaemia, infections disease and cancers. Compositions are useful

CC for treating disease associated with STAT6 activation and/or prevention

CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid

CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,

CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,

CC viral hepatitis and AIDS. The protein has efficient promoting STAT6

CC activity. The protein or nucleic acid is effectively useful for screening

CC compounds for treating and preventing disease associated with excessive

CC activation or inhibition of STAT6. The present sequence represents a

CC human cDNA encoding a protein which promotes STAT6 activation.

XX Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	20	Length:	2557
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-14 (1-9) x AD126160 (1-2557)

Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9

Db 1690 GTTTGTATTGACCGAAGGGGATA 1716

RESULT 36

AD126158

ID AD126158 standard; cDNA; 2557 BP.

XX AC AD126158;

XX 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #62.

ss; gene; human; signal transducer and activator of transcription 6;

STAT6; immunogen; STAT6 activation; allergy; inflammation;

autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;

Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;

systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

OS Homo sapiens.

XX WO2003104277-A2.

XX 18-DEC-2003.

XX 05-JUN-2003; 2003WO-JP007123.

XX 05-JUN-2002; 2002JP-00164257.

XX 06-JUN-2002; 2002US-0385912P.

XX 26-DEC-2002; 2002JP-00377326.

XX 27-DEC-2002; 2002US-0436467P.

XX 15-MAY-2003; 2003JP-00137505.

XX 16-MAY-2003; 2003US-0470836P.

XX

PA (ASAH) ASAH KASEI KK.
 XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX WPI; 2004-122214/12.
 DR P-PSDB; ADI26159.
 XX
 XX New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 XX Claim 4; SEQ ID NO 123; 1368pp; English.
 XX
 XX The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or an
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has effectively promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 20 Length: 2557
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-14 (1-9) x ADI26158 (1-2557)
 QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 DB 1690 GTTTTGATTTGACCGTAGGCGATA 1716
 RESULT 37
 ABA90338
 ID ABA90338 standard; cDNA; 1149 BP.
 XX
 AC ABA90338;
 XX
 XX 12-FEB-2002 (first entry)
 DT
 XX
 XX Human polynucleotide #13.
 KW Human; nootropic; neuroprotective; anticonvulsant; antidepressant;
 KW neuroleptic; tranquiliser; antiarrhythmic; cardiac; antiasthmatic;
 KW antiinflammatory; antihypertensive; hepatotropic; viricide; antidiabetic;
 KW nephrotropic; anorectic; cytostatic; vaccine; neurological disease;
 KW cardiovascular disease; respiratory disease; liver disease;
 KW renal disease; skeletal muscle disease; gastrointestinal disease;
 KW placental disease; testicular cancer; male fertility; pancreatic disease;
 KW ss.
 XX
 XX Homo sapiens.
 Os

XX WO200181363-A1.
 XX
 XX 01-NOV-2001.
 XX
 XX 26-APR-2001; 2001WO-US013360.
 XX
 XX 27-APR-2000; 2000US-0199963P.
 PR 11-MAY-2000; 2000US-020336P.
 PR 25-MAY-2000; 2000US-0207087P.
 PR 26-MAY-2000; 2000US-0207546P.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;
 PI Lai Y, Xie Q;
 XX
 XX WPI; 2002-0411392/05.
 DR P-PSDB; ABB53273.
 XX
 XX Novel polypeptides and polynucleotides useful as a vaccine for preventing
 PT and treating diseases associated the polypeptide, e.g. Alzheimer's
 PT disease, dyslipidemia, obesity, diabetes, infertility, asthma, amnesias.
 PT
 XX Claim 2; Page 51-52; 116pp; English.
 XX
 XX The invention relates to an isolated polypeptide comprising a 277, 480,
 CC 583, 581, 628, 424, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,
 CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 490, 462, 255,
 CC 784, 252, 593, 472, 607, 781, 640, 686 or 154 amino acid sequence as
 CC given in the specification. The polypeptides, modulators of the
 CC polypeptides and antibodies against the polypeptides are useful for
 CC treating diseases such as neurological and psychiatric diseases including
 CC Alzheimer's, paraeupranuclear palsy, Huntington's disease, myotonic
 CC dystrophy, anorexia and depression; cardiovascular diseases including
 CC congestive heart failure, Hodgkin's disease and myocardial infarction;
 CC respiratory diseases including asthma, chronic obstructive pulmonary
 CC disease, cystic fibrosis and adult respiratory distress syndrome; liver
 CC diseases including hypercholesterolaemia, cirrhosis, viral and nonviral
 CC hepatitis, Type II diabetes mellitus, and impaired glucose tolerance;
 CC renal disease including renal failure, acute tubular necrosis and
 CC glomerulonephritis; skeletal muscle diseases including Eulenburg's
 CC disease, hypoglycaemia and obesity; gastrointestinal diseases including
 CC myotonia congenita and intestinal obstruction; lymph diseases including
 CC lymphagiectasia; diseases of placenta including choriocarcinoma; diseases
 CC of testes including testicular cancer, male reproductive diseases
 CC including low testosterone and male infertility; and disease of pancreas
 CC including diabetic ketoacidosis, Type 1 and 2 diabetes and obesity. The
 CC present sequence encodes a polypeptide of the invention
 XX
 SQ Sequence 1149 BP; 120 A; 468 C; 417 G; 144 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 36 Length: 1149
 Score: 42.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 93.3% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-14 (1-9) x ABA90338 (1-1149)
 QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 DB 973 GTGCTTACCTAAACCGCGCGCATC 999
 RESULT 38
 ADU01816
 ID ADU01816 standard; cDNA; 3993 BP.
 XX
 AC ADU01816;
 XX

DT 27-JAN-2005 (first entry)
 XX Novel human polynucleotide seqid 283.
 XX
 XX cytostatic; antipsoriatic; antiinflammatory; gene therapy; Nanodisc;
 XX proliferative disorder; inflammatory disorder; immune disorder;
 XX metabolic disorder; bone disorder; CNS disorder; cancer; psoriasis;
 XX ulcerative colitis; human; gene; ss.
 XX
 XX Homo sapiens.
 XX
 XX WO2004093804-A2.
 XX
 XX 04-NOV-2004.
 XX
 XX 19-APR-2004; 2004WO-US012047.
 XX
 XX 18-APR-2003; 2003US-0463708P.
 XX
 XX 18-APR-2003; 2003US-0463732P.
 XX
 XX 02-MAY-2003; 2003US-0487199P.
 XX
 XX 02-MAY-2003; 2003US-0487230P.
 XX
 XX 19-MAY-2003; 2003US-0471306P.
 XX
 XX 19-MAY-2003; 2003US-0471336P.
 XX
 XX 08-JUL-2003; 2003US-0485223P.
 XX
 XX 08-JUL-2003; 2003US-0485224P.
 XX
 XX 14-JUL-2003; 2003US-0486466P.
 XX
 XX 14-JUL-2003; 2003US-0486480P.
 XX
 XX 08-AUG-2003; 2003US-0493573P.
 XX
 XX 08-AUG-2003; 2003US-0493577P.
 XX
 XX 08-SEP-2003; 2003US-0505059P.
 XX
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX
 XX Lee E, Hestir K, Chu K, Masuoka L, Williams LT;
 XX
 XX WPI; 2004-775861/76.
 XX
 XX P-PSDB; ADU02548.
 XX
 XX New first nucleic acid molecule comprising a polynucleotide sequence
 XX given in the specification, useful in preparing a composition for
 XX diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
 XX
 XX Claim 1; SEQ ID NO 283; 291pp; English.
 XX
 XX The invention describes a new first nucleic acid molecule comprising a
 XX polynucleotide sequence given in the specification. Also described are:
 XX an animal injected with the nucleic acid molecule; a second nucleic acid
 XX molecule comprising a second polynucleotide sequence that is at least
 XX about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or
 XX that hybridises to the first polynucleotide sequence under high
 XX stringency conditions; a vector comprising the nucleic acid molecule and
 XX a promoter that drives the expression of the nucleic acid molecule; a
 XX host cell transformed, transfected or infected with the nucleic acid molecule; a
 XX nucleic acid molecule; a nucleic acid composition comprising a carrier or
 XX a buffer and one or more compositions comprising the nucleic acid
 XX molecule, vector or host cell; a substantially purified polypeptide; an
 XX animal injected with the polypeptide; a polypeptide composition
 XX comprising the polypeptide molecule and a carrier or buffer; a cell
 XX culture medium comprising the polypeptide or transfected cells
 XX transfected with the polynucleotide; making a transformed, transfected,
 XX transduced, or infected host cell; synthesising Nanodiscs simultaneously
 XX and for synthesising a series of simultaneously-synthesised Nanodiscs
 XX sequentially utilising a dynamic system; preparing a hydrophobic protein
 XX for determination of crystal structure; immunising a non-human animal;
 XX screening for modulators of hydrophobic protein activity; a diagnostic
 XX kit; determining the presence of the nucleic acid molecule or its
 XX complement; determining the presence of an antibody to the polypeptide in
 XX a sample; an antibody specifically recognising, binding to or modulating
 XX the biological activity of at least one polypeptide encoded by a nucleic
 XX acid molecule or its biologically active fragment; an antibody
 XX composition comprising the antibody and a carrier; a bacteriophage, where
 XX the antibody is displayed on the bacteriophage; a bacterial cell
 XX comprising the bacteriophage; a non-human animal injected with the

CC antibody composition; a host cell that secretes the antibody; making an
 CC antibody; diagnosing a disease, disorder, syndrome, or condition
 CC comprising cancer, or proliferative, inflammatory, immune, metabolic,
 CC bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or
 CC conditions in a patient; a modulator composition comprising a modulator
 CC and a carrier; gene therapy; prophylactic or therapeutic treatment of a
 CC subject; an isolated modified cell comprising at least one first
 CC heterologous nucleic acid molecule, where the first heterologous nucleic
 CC acid molecule comprises a first polynucleotide sequence that encodes a
 CC first polypeptide; a non-human animal deficient in the polypeptide or
 CC that over-expresses the polypeptide; isolated tissues derived from the
 CC non-human animal; and one or more cells derived from the non-human
 CC animal. The nucleic acid is useful in preparing a composition for
 CC diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
 CC This sequence encodes a novel human polypeptide of the invention.
 XX
 XX Sequence 3993 BP; 719 A; 1298 C; 1224 G; 752 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 148 Length: 3993
 Score: 42.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 93.3% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-14 (1-9) x ADU01816 (1-3993)
 QY 1 ValLeuTyLeuAsnArgLysGlyIle 9
 |||||:|||||:|||||:|||||:|||||:
 Db 3820 GTGCTCTACTTAACCGCGCGGCATC 3846
 RESULT 39
 ABL11890
 ID ABL11890 standard; cDNA; 4289 BP.
 XX
 XX ABL11890;
 AC
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 30152.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW
 KW pharmaceutical; gene; ss.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 FN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 XX 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 FI
 XX WPI; 2001-656860/75.
 DR
 XX P-PSDB; ABB67787.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions.
 PT
 XX Claim 1; SEQ ID NO 30152; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (ABBS57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 4289 BP; 1219 A; 962 C; 967 G; 1141 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 160 Length: 4289
 Score: 42.00 Matches: 7
 Percent Similarity: 100.0% Conservatives: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 93.3% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-14 (1-9) x AB111890 (1-4289)

QY 1 ValLeuTyLeuAsnArgLysGlyIle 9
 : : : : :
 Db 3275 ATCATATATCTTAACCGTAAGGCATC 3301

RESULT 40
 AAS57114
 ID AAS57114 standard; DNA; 4289 BP.

XX AC AAS57114;

XX DT 16-JAN-2002 (first entry)

XX DE DNA encoding Drosophila G-protein coupled receptor, GPCR #22.

XX KW Drosophila; G-protein coupled receptor; GPCR; insecticide; diagnostic;
 KW mutation detection; ds.

XX OS Drosophila melanogaster.

XX PN WO200170980-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009341.

XX PR 23-MAR-2000; 2000US-0191638P.

XX PR 18-JUL-2000; 2000US-00618893.

XX PA (PEKE) PE CORP NY.

XX PI Cravchik A;

XX DR WPI; 2001-616405/71.

XX DR P-PSDB; AAU38944.

XX PT Sixty six Drosophila Melanogaster G-protein coupled receptors (GPCR),
 PT useful in the treatment and diagnosis of GPCR-related conditions and for
 PT identifying GPCR modulators for use as insecticides.
 XX Claim 4; Page 134-135; 392pp; English.

XX The invention relates to sixty six novel isolated Drosophila melanogaster
 CC G-protein coupled receptors (GPCR). The GPCR proteins and nucleic acids
 CC are useful in the treatment and diagnosis of GPCR-related conditions. The
 CC GPCR proteins and nucleic acids are also useful for identifying
 CC modulators of GPCR proteins for use as insecticides. The nucleic acid can
 CC also be used to detect mutations in GPCR genes and gene expression
 CC products such as mRNA. AAS57072-AAS57203 represent D. melanogaster G-
 CC coupled protein receptor genomic and coding sequences of the invention
 XX Sequence 4289 BP; 1219 A; 962 C; 967 G; 1141 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 160 Length: 4289

Score: 42.00 Matches: 7
 Percent Similarity: 100.0% Conservatives: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 93.3% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-14 (1-9) x AAS57114 (1-4289)

QY 1 ValLeuTyLeuAsnArgLysGlyIle 9
 : : : : :
 Db 3275 ATCATATATCTTAACCGTAAGGCATC 3301

RESULT 41

ADC35812

ID ADC35812 standard; DNA; 4289 BP.

XX AC ADC35812;

XX DT 18-DEC-2003 (first entry)

XX DE Drosophila G protein coupled receptor genomic DNA seq id 22.

XX KW G-protein coupled receptor; GPCR; insecticide; drug screening;
 KW insecticide screening; insecticidal activity; insecticidal tolerance;
 KW fruit fly; gene; ds.

XX OS Drosophila melanogaster.

XX PN US2003092124-A1.

XX PD 15-MAY-2003.

XX PF 15-OCT-2002; 2002US-00270333.

XX PR 03-DEC-1999; 98US-0168677P.

XX PR 12-JAN-2000; 2000US-0195691P.

XX PR 23-MAR-2000; 2000US-0191638P.

XX PR 18-JUL-2000; 2000US-00618893.

XX PA (APPL-) APPLERA CORP.

XX PI Cravchik A;

XX DR WPI; 2003-765480/72.

XX DR P-PSDB; ADC35814.

XX PT New isolated G-protein coupled receptor useful for identifying modulators
 PT as potential insecticides, to determine the biological activity of the
 PT protein and for identifying compounds that modulate receptor activity.

XX PS Claim 4; SEQ ID NO 64; 130pp; English.

XX The invention describes an isolated protein (I) consisting or comprising
 CC an amino acid sequence selected from fully defined 66 G-protein coupled
 CC receptor amino acid sequences (S1), as given in the specification, an
 CC allelic variant of (S1), an orthologue of (S1) or fragment of (S1). (I)
 CC is useful for identifying an agent that binds to (I) which comprises
 CC contacting the protein with an agent and assaying the contacted mixture
 CC to determine whether a complex is formed with the agent bound to the
 CC protein. (I) is useful for identifying modulators as potential
 CC insecticides, to determine the biological activity of the protein (a
 CC panel of multiple proteins for high-throughput screening), as targets for
 CC identifying agents for use in human drugs and for identifying compounds
 CC that modulate receptor activity. An antibody (II) that selectively binds
 CC to (I) is useful for assessing normal and aberrant subcellular
 CC localisation of cells and monitoring a treatment modality. A nucleic acid
 CC (III) encoding (I) is useful for drug/insecticide screening to identify
 CC compounds that modulate G-protein coupled receptor (GPCR) nucleic acid
 CC expression, diagnostic assays for qualitative changes in GPCR nucleic
 CC acid that lead to insecticidal activity/tolerance, to detect mutations in
 CC GPCR genes and gene expression products such as mRNA, and as
 CC hybridisation probes for determining the presence, level, form and
 CC distribution of nucleic acid expression. A host cell comprising a vector

CC containing (III) is useful for conducting cell-based assays involving the
CC GPCR protein or its fragments, and identifying GPCR protein mutants. This
CC sequence encodes a fruit fly G-protein coupled receptor (GPCR).
XX
SQ Sequence 4289 BP; 1219 A; 962 C; 967 G; 1141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 160 Length: 4289
Score: 42.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 93.3% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-14 (1-9) x ADC35812 (1-4289)

QY 1 ValLeuTyrLeuAenArgLysGlyIle 9
: : : : :
Db 3275 ATCATATATCTTAACCGTAAGGCATC 3301

RESULT 42

ADX27916
ID ADX27916 standard; cDNA; 874 BP.

XX
AC ADX27916;

XX
XX 21-APR-2005 (first entry)

XX
DE Plant full length insert polynucleotide seqid 10736.

XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.

XX
OS Unidentified.

XX
XX US2004034888-A1.

XX
XX 19-FEB-2004.

XX
XX 28-APR-2003; 2003US-00425114.

XX
XX 06-MAY-1999; 99US-00304517.

XX
XX 05-NOV-2001; 2001US-00985678.

XX
XX (LIUJ/) LIU J.

PA (ZHOU/) ZHOU Y.

PA (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.

PA (TAB/) TABASKA J E.

PA (CAOY/) CAO Y.

XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska J E, Cao Y;

XX
XX WPI; 2004-180133/17.

XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.

XX
PS Claim 1; SEQ ID NO 10736; 15pp; English.

XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp:seqdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as

CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.

XX
SQ Sequence 874 BP; 250 A; 160 C; 179 G; 285 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 43.5 Length: 874
Score: 41.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.1% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-14 (1-9) x ADX27916 (1-874)

QY 1 ValLeuTyrLeuAenArgLysGly 8
: : : : :
Db 712 GTTTGTATCTTAATAGAAAGGG 735

RESULT 43

ABD32791_2/c

Continuation (3 of 4) of ABD32791 from base 200001 (Human cancer-associated genomic DNA)

WP Sequence split into 4 fragments LOCUS ABD32791 Accession Abd32791

WP Fragment Name Begin End

WP ABD32791_0 1 110000

WP ABD32791_1 100001 210000

WP ABD32791_2 200001 310000

WP ABD32791_3 300001 350570

Alignment Scores:

Pred. No.: 1.05e+04 Length: 110000
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 91.1% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-14 (1-9) x ABD32791_2 (1-110000)

QY 1 ValLeuTyrLeuAenArgLysGlyIle 9
: : : : :
Db 48011 ATTCTGTATCTTAATAAAAGGGATT 47985

RESULT 44

ABD33549

ID ABD33549 standard; DNA; 99291 BP.

XX
AC ABD33549;

XX
XX 18-NOV-2004 (first entry)

XX
XX Human cancer-associated (CA) gene HD07-109.

XX
XX Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
ds; cancer; cytostatic.

XX
OS Homo sapiens.

XX
XX WO2004058146-A2.

XX
XX 15-JUL-2004.

XX

PF 15-DEC-2003; 2003WO-US040081.
 PR 17-DEC-2002; 2002US-00322281.
 XX (SAGR-) SAGRES DISCOVERY INC.
 XX Morris DW, Malandro MS;
 XX WPI; 2004-499109/47.
 XX Novel human cancer associated protein encoded within open reading frame
 PT of cancer associated gene, useful as targets for diagnosing cancer.
 XX Claim 16; SEQ ID NO 744; 182pp; English.
 PS The invention relates to cancer-associated proteins (CAP) and the cancer-
 CC associated (CA) nucleic acids encoding them. The invention also relates
 CC to a method for treating cancers involving administering to a patient an
 CC inhibitor of CAP, and a method of screening for anticancer activity in a
 CC potential drug involving providing a cell that expresses a CA gene,
 CC contacting a tissue sample derived from a cancer cell with an anticancer
 CC drug candidate and monitoring the effect of the anticancer drug candidate
 CC on expression of the CA gene. The CAP proteins are useful for detecting
 CC cancer associated with expression of a CAP protein in a test cell sample
 CC and for screening for a bioactive agent capable of modulating the
 CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 CC cancer, involving determining the expression of a CA nucleic acid in a
 CC tissue. This sequence represents a human CA gene of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 99291 BP; 26458 A; 20539 C; 22297 G; 29997 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.52e+04 Length: 99291
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 86.7% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-14 (1-9) x ABD33549 (1-99291)
 QY 2 LeuTyrLeuAsnArgLysGlyTle 9
 Db 30721 ATATATCTAAACAGAAAGGTATA 30744
 RESULT 45
 ADZ70593/c
 ID ADZ70593 standard; cDNA; 114596 BP.
 XX
 AC ADZ70593;
 XX
 DT 30-JUN-2005 (first entry)
 XX
 DE Human cDNA from lung cancer marker gene B4GALT5.
 XX
 KW Tumor marker; ss; gene; lung tumor; cytostatic; neoplasm; expression;
 KW DNA microarray.
 XX
 OS Homo sapiens.
 XX
 KW WO2005032495-A2.
 PN
 XX 14-APR-2005.
 PD
 XX
 XX 01-OCT-2004; 2004WO-US034163.
 PF
 XX 03-OCT-2003; 2003US-0508355P.
 PR
 XX (PAB) BAYER PHARM CORP.
 PA
 XX

PI Taylor I, Pauloski NR, Bigwood D;
 DR WPI; 2005-285325/29.
 DR P-PSDB; ADZ70594.
 XX
 XX Providing a patient diagnosis for lung cancer comprises comparing the
 PT level of expression of genes or gene products in a biological sample from
 PT the patient with that from a normal individual.
 XX
 XX Claim 2; SEQ ID NO 278; 60pp; English.
 PS
 XX The invention relates to providing a patient diagnosis for lung cancer
 CC comprising comparing the level of expression of genes or gene products in
 CC a biological sample from the patient with the level of expression of
 CC genes or gene products in a biological sample from a normal individual.
 CC Also included are distinguishing between normal and disease tissues,
 CC monitoring the response of a patient being treated for lung cancer by
 CC administering an anti-cancer agent, identifying a compound useful for the
 CC treatment of lung cancer and an array for distinguishing between normal
 CC and disease tissues (comprising 2 or more probes corresponding to 2 or
 CC more genes selected from any of the 200 nucleotide sequences given in the
 CC specification, or 2 or more polypeptides comprising any of the 200 amino
 CC acid sequences given in the specification). In providing a patient
 CC diagnosis for lung cancer, one or more genes are selected from any of the
 CC 200 nucleotide sequences as mentioned in the specification, or one or
 CC more gene products are polypeptides selected from any of the 20 amino
 CC acid sequences mentioned in the specification. The methods are useful for
 CC detecting and treating lung cancer. These may also be used for designing,
 CC identifying and optimizing therapeutics for cancer. The present sequences
 CC represents a cDNA from one of the 200 lung cancer marker genes. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 114596 BP; 33213 A; 26849 C; 24628 G; 29906 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.97e+04 Length: 114596
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 86.7% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-14 (1-9) x ADZ70593 (1-114596)
 QY 2 LeuTyrLeuAsnArgLysGlyTle 9
 Db 101082 ATATATCTAAACAGAAAGGTATA 101059
 RESULT 46
 ABK99972/c
 ID ABK99972 standard; DNA; 157875 BP.
 XX
 AC ABK99972;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human CADPKL genomic DNA.
 XX
 KW Human; calcium/calmodulin-dependent protein kinase-like gene; CADPKL; da;
 KW gene; neuropsychiatric disorder; attention deficit disorder; ADD;
 KW schizoaffective disorder; bipolar disorder; unipolar affective disorder;
 KW schizophrenia; adolescent conduct disorder; pharmacogenomics;
 KW fingerprinting; paternity testing; antidepressant; neuroleptic.
 XX
 OS Homo sapiens.
 XX
 XX WO200254939-A2.
 PN
 XX 18-JUL-2002.
 PD
 XX 07-JAN-2002; 2002WO-US000367.
 PF

XX 09-JAN-2001; 2001US-00757300.
 PR 23-AUG-2001; 2001US-00935464.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Meyer JM, Barrington-Martin R, Parker A;
 PI WPI; 2002-590643/63.
 XX
 DR New variants of calcium/calmodulin-dependent protein kinase-like nucleic
 XX acids and polypeptides, useful for diagnosing and treating
 PT neuropsychiatric disorders, e.g. schizophrenia, schizoaffective disorder,
 PT and bipolar disorder.
 XX
 PS Claim 1; Page 119-200; 223pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a polymorphic region
 CC of a Calcium/calmodulin-dependent protein kinase-like gene (CADPKL)
 CC allelic variant, and the polypeptide it encodes. CADPKL allelic variants
 CC are useful in determining whether a subject has or is at risk of
 CC developing a neuropsychiatric disorder, such as schizophrenia, attention
 CC deficit disorder (ADD), schizoaffective disorder, bipolar disorder,
 CC unipolar affective disorder and adolescent conduct disorder. The
 CC polypeptides, polynucleotides, antibodies and modulators of the CADPKL
 CC allelic variants are useful for diagnosing or treating these
 CC neuropsychiatric disorders. The polypeptides may be used to raise
 CC antibodies to a CADPKL polypeptide. The nucleic acids may be used as
 CC probes or primers, in pharmacogenomics for designing therapies for the
 CC disorders, and in fingerprinting for detection of different individuals
 CC with the same species (e.g. paternity testing). This sequence represents
 CC human CADPKL genomic DNA of the invention
 XX
 SQ Sequence 157875 BP; 44403 A; 35734 C; 34322 G; 43416 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.27e+04 Length: 157875
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 86.7% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-14 (1-9) x ABK99972 (1-157875)
 QY 2 LeuTyrLeuAsnArgLysGlyIle 9
 |||||:::|||||||
 DB 85262 CTTTATATGAACGGAAGGTATC 85239
 RESULT 47
 ABK75509
 ID ABK75509 standard; DNA; 1209 BP.
 AC ABK75509;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Bacillus licheniformis genomic sequence tag (GST) #2800.
 XX
 KW Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 OS Bacillus licheniformis.
 XX
 XX WO200229113-A2.
 XX
 XX 11-APR-2002.
 XX
 XX 05-OCT-2001; 2001WO-US031437.
 XX
 XX 06-OCT-2000; 2000US-00680598.
 PR 27-MAR-2001; 2001US-0279526P.
 PR

XX (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 XX Berka R, Clausen IG;
 XX WPI; 2002-416684/44.
 DR
 XX
 PT Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second Bacillus
 PT cells, by using substrate containing Bacillus genomic sequenced tag
 XX array.
 PS Claim 4; SEQ ID NO 2800; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive follow
 CC up characterisation is unnecessary, when one spot on an array equals one
 CC gene or one open reading frame, since sequence information is available.
 CC This sequence represents a genomic sequence tag (GST) used in the method
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1209 BP; 376 A; 289 C; 295 G; 249 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 281 Length: 1209
 Score: 38.00 Matches: 6
 Percent Similarity: 100.0% Conservative: 3
 Best Local Similarity: 66.7% Mismatches: 0
 Query Match: 84.4% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-14 (1-9) x ABK75509 (1-1209)
 QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
 ::|||:::|||||
 DB 674 ATCTGTACGTGAACAAAAAGGCATT 700
 RESULT 48
 ABQ33830
 ID ABQ33830 standard; DNA; 1859 BP.
 XX ABQ33830;
 AC ABQ33830;
 XX
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 20421.
 XX
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200218632-A2.
 XX
 XX 07-MAR-2002.
 PD

```

XX 01-SEP-2001; 2001WO-EP010074.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX 05-SEP-2000; 2000DE-01044543.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful for
XX diagnosis and prognosis, comprises selective hybridization of amplicons
XX from chemically treated DNA.
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one member,
XX of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
XX degree of hybridisation to both classes is determined from the label on
XX the amplicon. From the ratio of labels hybridised to the two classes of
XX oligomers, the degree of methylation is calculated. The method is used:
XX (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
XX and of a wide range of diseases, e.g. cancer, disorders of the central
XX nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
XX particularly by detecting mutations or single nucleotide polymorphisms
XX (SNP's); and (ii) for differentiation of cell or tissue types and for
XX investigating cell differentiation. The method allows the methylation
XX status of many C residues to be determined simultaneously. ABQ13410-
XX ABQ54121 represent genomic DNA sequences used to illustrate the method
XX for determining the degree of cytosine methylation described in the
XX disclosure of the invention
XX
XX Sequence 1859 BP; 243 A; 251 C; 648 G; 717 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 457 Length: 1859
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.4% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x ABQ33830 (1-1859)
QY 1 ValLeuTyrLeuAsnArgLysGlyTle 9
Db 1388 GTGTTTATTAAATCGTCGGGTATT 1414
RESULT 49
ABQ33831/c
ID ABQ33831 standard; DNA; 1859 BP.
XX
XX AC ABQ33831;
XX
XX 12-JUL-2002 (first entry)
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 20422.
XX
XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX drug; side effect; cancer; central nervous system; cardiovascular;
XX gastrointestinal; respiratory system; single nucleotide polymorphism;
XX SNP; cell differentiation; ds.
XX
XX Homo sapiens.
XX
XX WO200218632-A2.
XX
07-MAR-2002.
01-SEP-2001; 2001WO-EP010074.
01-SEP-2000; 2000DE-01043826.
05-SEP-2000; 2000DE-01044543.
(EPIG-) EPIGENOMICS AG.
Olek A, Piepenbrock C, Berlin K, Guetig D;
WPI; 2002-371829/40.
Determining the degree of cytosine methylation in genomic DNA, useful for
diagnosis and prognosis, comprises selective hybridization of amplicons
from chemically treated DNA.
Claim 12; 56pp + Sequence Listing; 56pp; German.
This invention describes a novel method for determining the degree of
methylation of a particular cytosine in a motif 5'-CpG-3', present in a
genomic sample of DNA. The sample is treated chemically to convert
cytosine (C) but not methylated C, to uracil, then part of the genomic
DNA that contains the target C is amplified to form a labeled amplicon.
The amplicon is hybridised to two classes, each with at least one member,
of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
degree of hybridisation to both classes is determined from the label on
the amplicon. From the ratio of labels hybridised to the two classes of
oligomers, the degree of methylation is calculated. The method is used:
(i) for diagnosis and/or prognosis of side effects of therapeutic drugs
and of a wide range of diseases, e.g. cancer, disorders of the central
nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
particularly by detecting mutations or single nucleotide polymorphisms
(SNP's); and (ii) for differentiation of cell or tissue types and for
investigating cell differentiation. The method allows the methylation
status of many C residues to be determined simultaneously. ABQ13410-
ABQ54121 represent genomic DNA sequences used to illustrate the method
for determining the degree of cytosine methylation described in the
disclosure of the invention
Sequence 1859 BP; 717 A; 648 C; 251 G; 243 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 457 Length: 1859
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.4% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x ABQ33831 (1-1859)
QY 1 ValLeuTyrLeuAsnArgLysGlyTle 9
Db 472 GTGTTTATTAAATCGTCGGGTATT 446
RESULT 50
ABQ71002/c
ID ABQ71002 standard; DNA; 3617 BP.
XX
XX AC ABQ71002;
XX
XX 29-AUG-2003 (revised)
XX 29-AUG-2002 (first entry)
XX
XX Listeria monocytogenes 4b contig DNA sequence #944.
XX
XX Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
XX Listeria monocytogenes ATCC 19115.
XX

```

```
PN WO200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR003061.
XX
PR 04-OCT-2000; 2000FR-00012697.
XX
PA (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI Kunst F, Glaser P;
XX
DR WPI; 2002-332479/37.
XX
PT New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators.
XX
PS Claim 14; SEQ ID NO 3815; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences (AB067188-
CC AB071212) from Listeria sp. The sequences are useful as probes and
CC primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of gene
CC expression. Proteins encoded by the nucleic acid sequences can be used to
CC screen for compounds that modulate gene expression, replication and
CC pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in anti-
CC Listeria vaccines. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
CC on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 3617 BP; 1098 A; 635 C; 837 G; 1043 T; 0 U; 4 Other;
```

Alignment Scores:

Pred. No.:	972	Length:	3617
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	87.5%	Mismatches:	0
Query Match:	84.4%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-14 (1-9) x AB071002 (1-3617)

Qy 2 LeuTyrLeuAsnArgLysGlyIle 9
Db 75 GTTACCTTAATCGAAAGGAATT 52

Search completed: April 25, 2006, 12:36:46
Job time : 336.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignments)
171.290 Million cell updates/sec

Title: US-10-774-176-14

Perfect score: 45

Sequence: 1 VYLNKRG1 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abs/ABSSWEB/spool/US10774176/runat_24042006_165114_19197/app_query.fasta_1
-DB=genEmbl -QFMT=fastcap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cgi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes04
-USER=US10774176 @CGN 1 1.6765 @runat_24042006_165114_19197 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sv.*

12: gb.un.*

13: gb.vi.*

14: gb.htg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	290	6	CQ687716 Sequence
2	45	100.0	421	10	Z94208 H.sapiens f
3	45	100.0	475	6	CQ920916 Sequence

4	45	100.0	901	6	BD249733	BD249733 Polypepti
5	45	100.0	901	6	AX025013	AX025013 Sequence
6	45	100.0	901	6	AX316088	AX316088 Sequence
7	45	100.0	927	6	AX829164	AX829164 Sequence
8	45	100.0	1260	6	AX467373	AX467373 Sequence
9	45	100.0	1260	6	AX821533	AX821533 Sequence
10	45	100.0	1260	6	AX821548	AX821548 Sequence
11	45	100.0	1263	6	BD249731	BD249731 Polypepti
12	45	100.0	1263	6	AX025011	AX025011 Sequence
13	45	100.0	1263	6	AX149553	AX149553 Sequence
14	45	100.0	1263	6	AX316086	AX316086 Sequence
15	45	100.0	1263	6	AX467371	AX467371 Sequence
16	45	100.0	1281	6	BD249732	BD249732 Polypepti
17	45	100.0	1281	6	AX025012	AX025012 Sequence
18	45	100.0	1281	6	AX316087	AX316087 Sequence
19	45	100.0	2053	8	HS5740A	229083 Homo sapien
20	45	100.0	2183	5	CR855786	CR855786 Xenopus t
21	45	100.0	2333	9	AF063939	AF063939 Rattus no
22	45	100.0	2359	6	BD127282	BD127282 Primer fo
23	45	100.0	2359	6	CQ782724	CQ782724 Sequence
24	45	100.0	2359	8	AK074786	AK074786 Homo sapi
25	45	100.0	2361	6	BD127283	BD127283 Primer fo
26	45	100.0	2361	6	CQ782726	CQ782726 Sequence
27	45	100.0	2361	6	AX961916	AX961916 Sequence
28	45	100.0	2361	8	AK074790	AK074790 Homo sapi
29	45	100.0	2361	9	BC087011	BC087011 Rattus no
30	45	100.0	2379	8	BC037161	BC037161 Homo sapi
31	45	100.0	2423	9	BC058198	BC058198 Mus muscu
32	45	100.0	2457	6	AX961912	AX961912 Sequence
33	45	100.0	2557	6	AX961914	AX961914 Sequence
34	45	100.0	2714	8	AB168308	AB168308 Macaca fa
35	45	100.0	5551	8	HS012159	HS012159 Homo sapi
36	45	100.0	7942	9	MMU012160	MMU012160 Mus muscu
37	45	100.0	121909	8	HSJ492P14	AL121977 Human DNA
38	45	100.0	167046	9	AC158516	AC158516 Mus muscu
39	45	100.0	210237	14	AC128294	AC128294 Rattus no
40	45	100.0	239076	14	AC106962	AC106962 Rattus no
41	43	95.6	163959	9	AC131704	AC131704 Mus muscu
42	42	93.3	615	6	CQ736619	CQ736619 Sequence
43	42	93.3	4289	6	CQ587318	CQ587318 Sequence
44	42	93.3	4289	6	AX254497	AX254497 Sequence
45	42	93.3	10066	14	AC018122	AC018122 Drosophil
46	42	93.3	64168	14	AC079003	AC079003 Homo sapi
47	42	93.3	111108	8	AF188024	AF188024 Homo sapi
48	42	93.3	114933	8	AC022363	AC022363 Homo sapi
49	42	93.3	169385	2	AC012164	AC012164 Drosophil
50	42	93.3	175837	14	AC022045	AC022045 Homo sapi
51	42	93.3	186200	9	AC111086	AC111086 Mus muscu
52	42	93.3	215225	8	AP001972	AP001972 Homo sapi
53	42	93.3	227050	14	AC099391	AC099391 Rattus no
54	42	93.3	228624	5	BX088524	BX088524 Zebrafish
55	42	93.3	245191	14	AC128815	AC128815 Rattus no
56	42	93.3	297236	2	AB003510	AB003510 Drosophil
57	41	91.1	556	10	BV244640	BV244640 S234P6291
58	41	91.1	57000	8	AP005020	AP005020 Homo sapi
59	41	91.1	57787	14	AC100680	AC100680 Mus muscu
60	41	91.1	64456	14	AC102915	AC102915 Mus muscu
61	41	91.1	77910	14	AP000452	AP000452 Homo sapi
62	41	91.1	88098	4	AC156374	AC156374 Loxodonta
63	41	91.1	121506	4	AC087807	AC087807 Felis cat
64	41	91.1	140771	9	AL591611	AL591611 Mouse DNA
65	41	91.1	153629	9	AC122426	AC122426 Mus muscu
66	41	91.1	173396	14	AC136059	AC136059 Rattus no
67	41	91.1	198667	9	AC111144	AC111144 Mus muscu
68	41	91.1	206584	14	AC153009	AC153009 Mus muscu
69	41	91.1	214975	14	AC123335	AC123335 Rattus no
70	41	91.1	222913	14	AC120325	AC120325 Rattus no
71	41	91.1	242405	14	AC145439	AC145439 Microcebu
72	41	91.1	246538	14	AC102978	AC102978 Rattus no
73	41	91.1	349980	6	CQ869999	CQ869999 Sequence
74	40	88.9	2263	15	AK118547	AK118547 Arabidops
75	40	88.9	100515	15	AC011665	AC011665 Arabidops
76	40	88.9	110000	15	AB017350_08	Continuation (9 of

C 77	40	88.9	116076	9	AL831718	AL931718	Mouse DNA	150	38	84.4	1101	1	AB024596	AB024596 Streptomy
C 78	40	88.9	124587	7	AP001346	AP001346	Homo sapi	151	38	84.4	1209	6	AX434385	AX434385 Sequence
C 79	40	88.9	127395	7	AY176327	AY176327	Staphyloc	152	38	84.4	1713	5	BC099984	BC099984 Danio rer
C 80	40	88.9	135513	8	BS000023	BS000023	Pan trogl	153	38	84.4	1765	5	AB097825	AB097825 Danio rer
C 81	40	88.9	138715	7	AY954969	AY954969	Bacterioph	154	38	84.4	2053	6	CQ731678	CQ731678 Sequence
C 82	40	88.9	156352	14	AC165193	AC165193	Eulemur m	155	38	84.4	2639	1	AB042601	AB042601 Streptomy
C 83	40	88.9	158593	14	AC141933	AC141933	Rattus no	156	38	84.4	2676	15	AB102671	AB102671 Streptomy
C 84	40	88.9	172727	9	AC106834	AC106834	Mus muscu	C 157	38	84.4	2960	1	AB042600	AB042600 Streptomy
C 85	40	88.9	178832	14	AC123979	AC123979	Rattus no	C 158	38	84.4	3617	6	AX416824	AX416824 Sequence
C 86	40	88.9	181094	14	AC152896	AC152896	Aotus nan	C 159	38	84.4	8509	6	AR271732	AR271732 Sequence
C 87	40	88.9	183268	8	AC123970	AC123970	Lemur cat	C 160	38	84.4	8509	6	AX281574	AX281574 Sequence
C 88	40	88.9	188829	8	AC146675	AC146675	Callithrix	C 161	38	84.4	8509	6	AX281868	AX281868 Sequence
C 89	40	88.9	196739	14	AC151369	AC151369	Aotus nan	C 162	38	84.4	8632	1	AB032065	AB032065 Streptomy
C 90	40	88.9	207519	9	AC153520	AC153520	Mus muscu	C 163	38	84.4	9090	1	AY392413	AY392413 Streptomy
C 91	40	88.9	226140	14	AC097415	AC097415	Rattus no	C 164	38	84.4	11714	1	AE013350	AE013350 Methanosa
C 92	40	88.9	233713	14	AC133034	AC133034	Rattus no	C 165	38	84.4	34548	8	HGA293565	HGA293565 Homo sapi
C 93	40	88.9	245390	9	AC103947	AC103947	Mus muscu	C 166	38	84.4	59704	14	AC017426	AC017426 Drosophil
C 94	40	88.9	264451	14	AC158594	AC158594	Mus muscu	C 167	38	84.4	70104	8	AL391828	AL391828 Human DNA
C 95	40	88.9	282915	14	AC098165	AC098165	Rattus no	C 168	38	84.4	72591	14	AC080137	AC080137 Homo sapi
C 96	40	88.9	340000	8	AP001666	AP001666	Homo sapi	C 169	38	84.4	75970	8	AC096668	AC096668 Homo sapi
C 97	39	86.7	42353	8	AL158156	AL158156	Human DNA	C 170	38	84.4	87000	8	AC097103	AC097103 Homo sapi
C 98	39	86.7	53047	14	AC153576	AC153576	Rattus no	C 171	38	84.4	90877	8	AC069436	AC069436 Homo sapi
C 99	39	86.7	57272	14	AC100158	AC100158	Mus muscu	C 172	38	84.4	92852	9	AL671489	AL671489 Mouse DNA
C 100	39	86.7	93292	5	BX928746	BX928746	Zebrafish	C 173	38	84.4	92852	9	AL671489	AL671489 Mouse DNA
C 101	39	86.7	97375	8	AL353898	AL353898	Human DNA	C 174	38	84.4	110000	1	AB017333_11	AB017333_11 Continuation (12 o
C 102	39	86.7	101335	9	AL928822	AL928822	Mouse DNA	C 175	38	84.4	110000	1	CP000002_11	CP000002_11 Continuation (12 o
C 103	39	86.7	113368	8	AC123965	AC123965	Macaca mu	C 176	38	84.4	110000	14	AC011600_1	AC011600_1 Continuation (2 of
C 104	39	86.7	114596	8	HS1063B2	HS1063B2	Human DNA	C 177	38	84.4	110000	15	AP008213_000	AP008213_000 Oryza sat
C 105	39	86.7	120187	8	AC005884	AC005884	Homo sapi	C 178	38	84.4	110159	14	AC068178	AC068178 Homo sapi
C 106	39	86.7	127735	9	AC159615	AC159615	Mus muscu	C 179	38	84.4	110931	14	AC027079	AC027079 Mus muscu
C 107	39	86.7	152080	5	AL953877	AL953877	Zebrafish	C 180	38	84.4	111402	9	AP003154	AP003154 Mus muscu
C 108	39	86.7	153878	14	AC159170	AC159170	Papio anu	C 181	38	84.4	112924	9	AC025500	AC025500 Mus muscu
C 109	39	86.7	157875	8	H82721L6	H82721L6	Human DNA	C 182	38	84.4	123203	8	AC003036	AC003036 Homo sapi
C 110	39	86.7	163349	9	AL672233	AL672233	Mouse DNA	C 183	38	84.4	125662	14	AC153099	AC153099 Ginglymos
C 111	39	86.7	168929	8	AC084730	AC084730	Papio anu	C 184	38	84.4	129938	8	AC008071	AC008071 Homo sapi
C 112	39	86.7	171076	14	AC154891	AC154891	Bos tauru	C 185	38	84.4	131386	14	AL355514	AL355514 Homo sapi
C 113	39	86.7	174565	14	AC142557	AC142557	Cercopith	C 186	38	84.4	132050	8	AC004909	AC004909 Homo sapi
C 114	39	86.7	179523	14	AC156897	AC156897	Bos tauru	C 187	38	84.4	132532	8	AC142341	AC142341 Pan trogl
C 115	39	86.7	180574	8	AC092807	AC092807	Homo sapi	C 188	38	84.4	139968	8	AL356140	AL356140 Human DNA
C 116	39	86.7	182616	9	AC133494	AC133494	Mus muscu	C 189	38	84.4	142227	14	AL162492	AL162492 Homo sapi
C 117	39	86.7	188632	14	CR392351	CR392351	Danio rer	C 190	38	84.4	143890	8	HS6628N3	HS6628N3 Human DNA
C 118	39	86.7	191105	14	AC150473	AC150473	Papio anu	C 191	38	84.4	148183	9	AL645591	AL645591 Mouse DNA
C 119	39	86.7	192856	14	AC155623	AC155623	Zea mays	C 192	38	84.4	149572	8	AC004695	AC004695 Homo sapi
C 120	39	86.7	193560	14	AC123261	AC123261	Mus muscu	C 193	38	84.4	150519	5	CR847826	CR847826 Zebrafish
C 121	39	86.7	199504	14	AC164925	AC164925	Colobus g	C 194	38	84.4	151766	8	AC016396	AC016396 Homo sapi
C 122	39	86.7	200238	15	AC149290	AC149290	Solanum d	C 195	38	84.4	152666	9	AC131765	AC131765 Mus muscu
C 123	39	86.7	202971	14	CR759948	CR759948	Danio rer	C 196	38	84.4	154954	14	AC161739	AC161739 Dasyypus n
C 124	39	86.7	208707	8	AC084729	AC084729	Papio anu	C 197	38	84.4	155156	8	AL590682	AL590682 Human DNA
C 125	39	86.7	209811	14	CT009723	CT009723	Mus muscu	C 198	38	84.4	156444	9	AP003153	AP003153 Mus muscu
C 126	39	86.7	210476	9	AC164045	AC164045	Bos tauru	C 199	38	84.4	157633	14	AC137826	AC137826 Medicago
C 127	39	86.7	210950	14	AC129317	AC129317	Mus muscu	C 200	38	84.4	157633	14	AC137826	AC137826 Medicago
C 128	39	86.7	214562	14	AC021628	AC021628	Mus muscu	C 201	38	84.4	157843	14	AC153712	AC153712 Bos tauru
C 129	39	86.7	223468	14	AC152880	AC152880	Bos tauru	C 202	38	84.4	160655	8	AC064874	AC064874 Homo sapi
C 130	39	86.7	230185	14	AC021579	AC021579	Mus muscu	C 203	38	84.4	164611	8	AC068675	AC068675 Homo sapi
C 131	39	86.7	231260	14	AL160172	AL160172	Homo sapi	C 204	38	84.4	165496	9	AC165271	AC165271 Mus muscu
C 132	39	86.7	233046	14	AC162249	AC162249	Bos tauru	C 205	38	84.4	167315	14	AC152375	AC152375 Dasyypus n
C 133	39	86.7	242407	14	AC155470	AC155470	Zea mays	C 206	38	84.4	168364	5	AC147788	AC147788 Latimeria
C 134	39	86.7	300451	14	AC158046	AC158046	Bos tauru	C 207	38	84.4	170240	15	AP005869	AP005869 Oryza sat
C 135	39	86.7	325076	14	AC125912	AC125912	Rattus no	C 208	38	84.4	172615	14	AC139765	AC139765 Homo sapi
C 136	38	84.4	550	15	AF365161	AF365161	Hymenaea	C 209	38	84.4	173422	2	AC007467	AC007467 Drosophil
C 137	38	84.4	558	15	AF365162	AF365162	Hymenaea	C 210	38	84.4	173428	8	AC069575	AC069575 Homo sapi
C 138	38	84.4	591	15	AF365160	AF365160	Hymenaea	C 211	38	84.4	173509	8	AP002436	AP002436 Homo sapi
C 139	38	84.4	5920	2	AY722997	AY722997	Pamborus	C 212	38	84.4	174190	9	AC125518	AC125518 Mus muscu
C 140	38	84.4	1020	2	AY722998	AY722998	Pamborus	C 213	38	84.4	176053	14	AC068190	AC068190 Homo sapi
C 141	38	84.4	1020	2	AY722999	AY722999	Pamborus	C 214	38	84.4	178456	8	AP003170	AP003170 Homo sapi
C 142	38	84.4	1020	2	AY723000	AY723000	Pamborus	C 215	38	84.4	180337	9	AC158962	AC158962 Mus muscu
C 143	38	84.4	1020	2	AY723001	AY723001	Pamborus	C 216	38	84.4	183029	14	AC022067	AC022067 Homo sapi
C 144	38	84.4	1020	2	AY723018	AY723018	Pamborus	C 217	38	84.4	185975	9	AC101852	AC101852 Mus muscu
C 145	38	84.4	1020	2	AY723019	AY723019	Pamborus	C 218	38	84.4	187607	5	AC150284	AC150284 Latimeria
C 146	38	84.4	1020	2	AY723020	AY723020	Pamborus	C 219	38	84.4	187995	9	AC127254	AC127254 Mus muscu
C 147	38	84.4	1020	2	AY723021	AY723021	Pamborus	C 220	38	84.4	189876	8	AL357892	AL357892 Human DNA
C 148	38	84.4	1020	2	AY723022	AY723022	Pamborus	C 221	38	84.4	190489	8	AC067794	AC067794 Homo sapi
C 149	38	84.4	1020	2	AY723023	AY723023	Pamborus	C 222	38	84.4	190960	14	AC036188	AC036188 Homo sapi

C 369	37	82.2	137230	8	AC027793	AC027793 Homo sapi	442	37	82.2	204608	9	AC122281	AC122281 Mus muscu
C 370	37	82.2	140454	15	OSJN00183	AL662984 Oryza sat	443	37	82.2	205800	5	EX072534	EX072534 Zebrafish
C 371	37	82.2	141415	8	AC087214	AC087214 Papio anu	444	37	82.2	207872	14	AC163883	AC163883 Bos tauru
C 372	37	82.2	143322	9	AC115440	AC115440 Rattus no	445	37	82.2	208407	9	AC163014	AC163014 Mus muscu
C 373	37	82.2	144422	5	BX294389	BX294389 Zebrafish	446	37	82.2	208693	14	AC160038	AC160038 Bos tauru
C 374	37	82.2	144082	8	AL139115	AL139115 Human DNA	C 447	37	82.2	210752	14	AC098606	AC098606 Rattus no
C 375	37	82.2	150159	4	AB192510	AB192510 Sus scrofa	C 448	37	82.2	210972	14	AC090288	AC090288 Mus muscu
C 376	37	82.2	150704	14	AC026721	AC026721 Homo sapi	C 449	37	82.2	211091	9	AC012540	AC012540 Mus muscu
C 377	37	82.2	151008	14	AC035146	AC035146 Homo sapi	450	37	82.2	211680	9	AL603682	AL603682 Mouse DNA
C 378	37	82.2	152036	14	AL161776	AL161776 Homo sapi	451	37	82.2	212934	8	AC138207	AC138207 Homo sapi
C 379	37	82.2	153791	14	AC119629	AC119629 Rattus no	452	37	82.2	213068	14	AC156348	AC156348 Colobus g
C 380	37	82.2	156075	14	CR848822	CR848822 Danio rer	C 453	37	82.2	214196	9	AC116675	AC116675 Mus muscu
C 381	37	82.2	158500	9	AC131983	AC131983 Mus muscu	C 454	37	82.2	214485	14	AC123416	AC123416 Rattus no
C 382	37	82.2	159206	15	AP005826	AP005826 Oryza sat	455	37	82.2	215622	5	BX465862	BX465862 Zebrafish
C 383	37	82.2	160312	8	AC010626	AC010626 Homo sapi	456	37	82.2	215250	14	AC116190	AC116190 Rattus no
C 384	37	82.2	160587	8	AL158064	AL158064 Human DNA	457	37	82.2	215892	14	AC152260	AC152260 Bos tauru
C 385	37	82.2	160850	8	AC130184	AC130184 Macaca mu	C 458	37	82.2	216267	14	AC152848	AC152848 Callithri
C 386	37	82.2	160875	14	CR786575	CR786575 Danio rer	459	37	82.2	216748	14	AC114177	AC114177 Rattus no
C 387	37	82.2	161004	15	AP003255	AP003255 Oryza sat	460	37	82.2	217514	5	CR847503	CR847503 Zebrafish
C 388	37	82.2	161153	8	AC093901	AC093901 Homo sapi	461	37	82.2	218884	14	AC161536	AC161536 Mus muscu
C 389	37	82.2	161266	9	AC127298	AC127298 Mus muscu	462	37	82.2	220333	14	AC158141	AC158141 Mus muscu
C 390	37	82.2	161581	14	AC149630	AC149630 Rhinoloph	463	37	82.2	221729	14	AC094575	AC094575 Rattus no
C 391	37	82.2	161614	2	AC011706	AC011706 Drosophila	C 464	37	82.2	223408	8	AC005610	AC005610 Homo sapi
C 392	37	82.2	163443	14	AC006280	AC006280 Plasmodiu	C 465	37	82.2	223797	14	AC021978	AC021978 Homo sapi
C 393	37	82.2	163632	14	AC069057	AC069057 Homo sapi	C 466	37	82.2	224373	14	AC098129	AC098129 Rattus no
C 394	37	82.2	164435	14	AC153073	AC153073 Cercopith	C 467	37	82.2	224500	14	AC136530	AC136530 Rattus no
C 395	37	82.2	165555	14	AC164635	AC164635 Mus muscu	468	37	82.2	224613	14	BX901945	BX901945 Danio rer
C 396	37	82.2	167097	2	AC023747	AC023747 Drosophila	C 469	37	82.2	224948	14	AC135884	AC135884 Rattus no
C 397	37	82.2	167743	9	AC140259	AC140259 Mus muscu	470	37	82.2	225277	14	AC129863	AC129863 Rattus no
C 398	37	82.2	168556	5	BX842240	BX842240 Zebrafish	C 471	37	82.2	227533	9	AC102856	AC102856 Mus muscu
C 399	37	82.2	169070	9	AL845299	AL845299 Mouse DNA	C 472	37	82.2	227632	8	AC087382	AC087382 Homo sapi
C 400	37	82.2	171090	14	AC137166	AC137166 Rattus no	473	37	82.2	227936	14	AC119370	AC119370 Rattus no
C 401	37	82.2	171523	8	AC010206	AC010206 Homo sapi	474	37	82.2	229632	9	AC107667	AC107667 Mus muscu
C 402	37	82.2	171944	14	AC026531	AC026531 Homo sapi	C 475	37	82.2	231914	14	CR752646	CR752646 Danio rer
C 403	37	82.2	171944	14	AC138724	AC138724 Cercopith	476	37	82.2	233344	14	AC097829	AC097829 Rattus no
C 404	37	82.2	172778	5	CR788302	CR788302 Zebrafish	C 477	37	82.2	242269	14	AC133719	AC133719 Rattus no
C 405	37	82.2	172876	8	AC114489	AC114489 Homo sapi	C 478	37	82.2	242351	14	AC162468	AC162468 Bos tauru
C 406	37	82.2	172953	8	AC067721	AC067721 Homo sapi	479	37	82.2	243177	14	AC137437	AC137437 Rattus no
C 407	37	82.2	173073	15	AC139170	AC139170 Oryza sat	C 480	37	82.2	243232	14	AC096462	AC096462 Rattus no
C 408	37	82.2	173414	8	AC010387	AC010387 Homo sapi	C 481	37	82.2	244402	14	AC098222	AC098222 Rattus no
C 409	37	82.2	174162	14	AC027409	AC027409 Homo sapi	482	37	82.2	244662	14	AC111927	AC111927 Rattus no
C 410	37	82.2	174239	14	BX682232	BX682232 Mus muscu	483	37	82.2	245136	14	AC111706	AC111706 Rattus no
C 411	37	82.2	174580	14	AC022649	AC022649 Homo sapi	C 484	37	82.2	247161	9	AC162856	AC162856 Mus muscu
C 412	37	82.2	174820	2	AC010917	AC010917 Drosophila	C 485	37	82.2	248722	14	AC095781	AC095781 Rattus no
C 413	37	82.2	174992	14	AC025422	AC025422 Homo sapi	486	37	82.2	248871	14	AC094352	AC094352 Rattus no
C 414	37	82.2	176054	9	AC122214	AC122214 Mus muscu	C 487	37	82.2	248895	14	AC115311	AC115311 Rattus no
C 415	37	82.2	176580	8	BS000178	BS000178 Pan trogl	488	37	82.2	249487	9	MM025178	MM025178 Mus muscu
C 416	37	82.2	177717	8	AC061958	AC061958 Homo sapi	489	37	82.2	249487	9	MM025183	MM025183 Mus muscu
C 417	37	82.2	178026	9	AC102922	AC102922 Mus muscu	C 490	37	82.2	250384	14	AC102955	AC102955 Rattus no
C 418	37	82.2	178057	9	AC125097	AC125097 Mus muscu	C 491	37	82.2	251075	14	CR931806	CR931806 Danio rer
C 419	37	82.2	179597	14	CR854892	CR854892 Danio rer	492	37	82.2	251251	14	AC120734	AC120734 Rattus no
C 420	37	82.2	182101	9	AC133935	AC133935 Mus muscu	493	37	82.2	253132	2	AE014846	AE014846 Plasmodiu
C 421	37	82.2	182334	14	AC166601	AC166601 Nomasus	494	37	82.2	255144	14	AC130901	AC130901 Rattus no
C 422	37	82.2	183494	14	AC079989	AC079989 Rattus no	495	37	82.2	255956	9	AC110237	AC110237 Mus muscu
C 423	37	82.2	184149	9	AC128862	AC128862 Rattus no	C 496	37	82.2	259105	14	AL591066	AL591066 Homo sapi
C 424	37	82.2	184333	5	BX470078	BX470078 Zebrafish	497	37	82.2	260082	14	AC130175	AC130175 Rattus no
C 425	37	82.2	184352	8	BS000179	BS000179 Pan trogl	C 498	37	82.2	260336	14	AC094614	AC094614 Rattus no
C 426	37	82.2	184424	14	AC106319	AC106319 Rattus no	C 499	37	82.2	268510	14	AC095840	AC095840 Rattus no
C 427	37	82.2	185419	8	AC016113	AC016113 Homo sapi	C 500	37	82.2	278172	9	AC115725	AC115725 Mus muscu
C 428	37	82.2	185531	14	AC0012059	AC0012059 Homo sapi	501	37	82.2	287560	1	AE017274	AE017274 Bacillus
C 429	37	82.2	185566	8	AC004168	AC004168 Homo sapi	C 502	37	82.2	288610	14	AC120486	AC120486 Rattus no
C 430	37	82.2	187490	8	AC023902	AC023902 Homo sapi	C 503	37	82.2	290973	14	AC093944	AC093944 Rattus no
C 431	37	82.2	187601	15	OSJN00079	AL606645 Oryza sat	504	37	82.2	300051	2	AE003370	AE003370 Drosophila
C 432	37	82.2	188637	9	AC108915	AC108915 Mus muscu	505	37	82.2	310993	2	AE003431	AE003431 Drosophila
C 433	37	82.2	188836	9	AC102606	AC102606 Mus muscu	C 506	37	82.2	329234	14	AC163334	AC163334 Mus muscu
C 434	37	82.2	189137	9	AC079378	AC079378 Rattus no	C 507	37	82.2	347664	9	BX883043	BX883043 Rattus no
C 435	37	82.2	191607	9	AC123599	AC123599 Mus muscu	508	37	82.2	348946	9	BX883051	BX883051 Rattus no
C 436	37	82.2	192589	8	AC093157	AC093157 Homo sapi	509	36	80.0	263	2	AF467546	AF467546 Diacamma
C 437	37	82.2	196764	9	AC153012	AC153012 Mus muscu	510	36	80.0	263	2	AF467547	AF467547 Diacamma
C 438	37	82.2	197708	14	AC165380	AC165380 Colobus g	511	36	80.0	263	2	AF467548	AF467548 Diacamma
C 439	37	82.2	197958	14	AC098366	AC098366 Rattus no	512	36	80.0	263	2	AF467549	AF467549 Diacamma
C 440	37	82.2	198927	14	AC026863	AC026863 Homo sapi	513	36	80.0	263	2	AF467550	AF467550 Diacamma
C 441	37	82.2	199359	14	CR925762	CR925762 Danio rer	514	36	80.0	263	2	AF467551	AF467551 Diacamma

515	36	80.0	263	2	AP467552	AP467552 Diacamma	588	36	80.0	63610	8	AC092990	AC092990 Homo sapi
516	36	80.0	263	2	AP467553	AP467553 Diacamma	C 589	36	80.0	65177	15	AP001313	AP001313 Arabidops
517	36	80.0	273	6	AR450937	AR450937 Sequence	C 590	36	80.0	66850	14	AC068332	AC068332 Homo sapi
518	36	80.0	518	10	BV272201	BV272201 S23P6143	C 591	36	80.0	68927	14	AC136854	AC136854 Rattus no
C 519	36	80.0	563	10	BV425715	BV425715 S23P6313	C 592	36	80.0	71549	8	AC008476	AC008476 Homo sapi
C 520	36	80.0	583	10	BV341745	BV341745 S23P66112	C 593	36	80.0	75234	15	AP004941	AP004941 Lotus cor
C 521	36	80.0	591	10	BV320398	BV320398 S23P66439	C 594	36	80.0	75678	14	AC093372	AC093372 Mus muscu
C 522	36	80.0	600	10	BV319630	BV319630 S23P66355	C 595	36	80.0	81971	8	HS593C16	HS593C16 Mus muscu
523	36	80.0	604	10	BV429550	BV429550 S23P6472	C 596	36	80.0	83989	8	AL603713	AL603713 Human DNA
524	36	80.0	616	2	AGA237664	AGA237664 Anopheles	C 597	36	80.0	84157	14	AC016260	AC016260 Homo sapi
525	36	80.0	619	15	EGU271979	EGU271979 Elaeis gu	C 598	36	80.0	84633	14	AC144966	AC144966 Xenopus t
526	36	80.0	669	10	BV225675	BV225675 S23P6183	C 599	36	80.0	85189	14	AC166588	AC166588 Bos tauru
527	36	80.0	719	10	BV656761	BV656761 S21P6166	C 600	36	80.0	85206	14	AC159459	AC159459 Pan trogl
C 528	36	80.0	739	8	HS344245	HS344245 Homo sapi	C 601	36	80.0	85579	8	AC004744	AC004744 Homo sapi
C 529	36	80.0	799	10	BV599117	BV599117 S21P6178	C 602	36	80.0	91084	5	AC137997	AC137997 Oryza sat
C 530	36	80.0	850	15	AY504786	AY504786 Arcototie	C 603	36	80.0	93780	5	CR936411	CR936411 Zebrafish
C 531	36	80.0	852	15	AY504787	AY504787 Arcototie	C 604	36	80.0	94444	4	AC090961	AC090961 Bos tauru
C 532	36	80.0	872	10	BV479423	BV479423 sqi95a03	C 605	36	80.0	96426	14	AC014908	AC014908 Drosophil
C 533	36	80.0	1069	2	AP231694	AP231694 Carabus c	C 606	36	80.0	103615	15	AC105767	AC105767 Oryza sat
C 534	36	80.0	1083	2	AB092699	AB092699 Carabus c	C 607	36	80.0	105230	14	AL732359	AL732359 Continuation (13 o
C 535	36	80.0	1083	2	AB092700	AB092700 Carabus c	C 608	36	80.0	105350	14	CR936407	CR936407 Panio rer
C 536	36	80.0	1083	2	AB092701	AB092701 Carabus c	C 609	36	80.0	106301	8	AC008602	AC008602 Homo sapi
C 537	36	80.0	1083	2	AB092702	AB092702 Carabus c	C 610	36	80.0	107640	8	AC084852	AC084852 Homo sapi
C 538	36	80.0	1083	2	AB092703	AB092703 Carabus c	C 611	36	80.0	108879	15	ATT12K4	ATT12K4 Arabidops
C 539	36	80.0	1083	2	AB092704	AB092704 Carabus c	C 612	36	80.0	109156	8	AB065679	AB065679 Homo sapi
C 540	36	80.0	1083	2	CCOMTNDSSB	CCOMTNDSSB Carabus (Au	C 613	36	80.0	109442	1	CP000031	CP000031 Continuation (41 o
C 541	36	80.0	1084	2	AB101025	AB101025 Carabus b	C 614	36	80.0	109720	5	AL714031	AL714031 Zebrafish
C 542	36	80.0	1084	2	AB101033	AB101033 Carabus a	C 615	36	80.0	110000	1	CR767821	CR767821 Continuation (12 o
C 543	36	80.0	1084	2	AB101034	AB101034 Carabus a	C 616	36	80.0	110000	1	CR925677	CR925677 Continuation (12 o
C 544	36	80.0	1281	8	AY823398	AY823398 Homo sapi	C 617	36	80.0	110000	1	CR925678	CR925678 Continuation (12 o
C 545	36	80.0	1349	6	AR354593	AR354593 Sequence	C 618	36	80.0	110000	1	AB014291	AB014291 Continuation (16 o
C 546	36	80.0	1349	6	AR356149	AR356149 Sequence	C 619	36	80.0	110000	1	AB017223	AB017223 Continuation (17 o
C 547	36	80.0	1385	6	BD217648	BD217648 Human SOC	C 620	36	80.0	110000	1	BA000004	BA000004 Continuation (23 o
C 548	36	80.0	1468	15	BT018388	BT018388 Zean mayas	C 621	36	80.0	110000	1	CP000046	CP000046 Continuation (20 o
C 549	36	80.0	1750	15	AK071183	AK071183 Oryza sat	C 622	36	80.0	110000	2	CP000046	CP000046 Continuation (21 o
C 550	36	80.0	2095	6	AK713859	AK713859 Sequence	C 623	36	80.0	110000	2	AB003524	AB003524 Drosophil
C 551	36	80.0	2095	8	AK055741	AK055741 Homo sapi	C 624	36	80.0	110000	6	BD430793	BD430793 Continuation (14 o
C 552	36	80.0	2315	6	CQ605819	CQ605819 Sequence	C 625	36	80.0	110000	6	BD430793	BD430793 Continuation (15 o
C 553	36	80.0	2361	1	AJ872071	AJ872071 Bacillus	C 626	36	80.0	110000	14	AC091229	AC091229 Continuation (6 of
C 554	36	80.0	2361	6	AX924206	AX924206 Sequence	C 627	36	80.0	110000	14	AC091347	AC091347 Continuation (3 of
C 555	36	80.0	2435	5	AF028805	AF028805 Xenopus l	C 628	36	80.0	110000	14	AC091347	AC091347 Continuation (4 of
C 556	36	80.0	2632	8	BC028696	BC028696 Homo sapi	C 629	36	80.0	110000	14	AC108583	AC108583 Continuation (2 of
C 557	36	80.0	3375	6	AX695700	AX695700 Sequence	C 630	36	80.0	110000	15	AP008214	AP008214 Continuation (169
C 558	36	80.0	3642	14	AC019549	AC019549 Drosophil	C 631	36	80.0	110000	15	AP008214	AP008214 Continuation (170
C 559	36	80.0	3728	6	AX695699	AX695699 Sequence	C 632	36	80.0	110000	15	AP008214	AP008214 Continuation (258
C 560	36	80.0	3728	9	MM07012	MM07012 Mus muscu	C 633	36	80.0	110000	15	AP016818	AP016818 Continuation (11 o
C 561	36	80.0	4331	14	AC014865	AC014865 Drosophil	C 634	36	80.0	110000	15	AB016818	AB016818 Continuation (12 o
C 562	36	80.0	4696	15	AK120094	AK120094 Oryza sat	C 635	36	80.0	110000	15	AP008207	AP008207 Continuation (31 o
C 563	36	80.0	5494	15	AK121450	AK121450 Oryza sat	C 636	36	80.0	110000	15	AP008207	AP008207 Continuation (319
C 564	36	80.0	10930	1	AB009480	AB009480 Brucella	C 637	36	80.0	110000	15	AP008209	AP008209 Continuation (293
C 565	36	80.0	13650	1	AF285969	AF285969 Salmonell	C 638	36	80.0	110000	15	AP008210	AP008210 Continuation (79 o
C 566	36	80.0	13716	1	AB014121	AB014121 Buchnera	C 639	36	80.0	110000	15	AP008210	AP008210 Continuation (80 o
C 567	36	80.0	16463	6	CQ578585	CQ578585 Sequence	C 640	36	80.0	110000	15	AP008211	AP008211 Continuation (126
C 568	36	80.0	18425	14	AC012734	AC012734 Drosophil	C 641	36	80.0	110000	15	AP008211	AP008211 Continuation (182
C 569	36	80.0	20113	14	AC015413	AC015413 Drosophil	C 642	36	80.0	110000	15	AP008211	AP008211 Continuation (183
C 570	36	80.0	29485	2	AC005439	AC005439 Drosophil	C 643	36	80.0	110000	15	AP008212	AP008212 Continuation (223
C 571	36	80.0	31697	8	HS39865	HS39865 Human DNA s	C 644	36	80.0	110952	14	AP004042	AP004042 Oryza sat
C 572	36	80.0	34859	14	AC020483	AC020483 Drosophil	C 645	36	80.0	112031	8	AL445487	AL445487 Human DNA
C 573	36	80.0	36069	1	CR377166	CR377166 Circul	C 646	36	80.0	112115	8	AC093208	AC093208 Homo sapi
C 574	36	80.0	36069	6	CQ972254	CQ972254 Sequence	C 647	36	80.0	114290	8	AL138814	AL138814 Human DNA
C 575	36	80.0	38174	2	AF099915	AF099915 Caenorhab	C 648	36	80.0	117636	8	HSJ365012	HSJ365012 Human DNA
C 576	36	80.0	40111	8	AC005266	AC005266 Homo sapi	C 649	36	80.0	117995	9	AL929179	AL929179 Mouse DNA
C 577	36	80.0	40229	8	AC074139	AC074139 Homo sapi	C 650	36	80.0	118616	8	AL157706	AL157706 Human DNA
C 578	36	80.0	45687	8	AC134730	AC134730 Homo sapi	C 651	36	80.0	121008	8	AC078917	AC078917 Homo sapi
C 579	36	80.0	50954	8	HSJ272H18	HSJ272H18 Human DNA	C 652	36	80.0	123070	8	AC020656	AC020656 Homo sapi
C 580	36	80.0	55835	14	AC091645	AC091645 Homo sapi	C 653	36	80.0	124695	15	AP003913	AP003913 Oryza sat
C 581	36	80.0	56168	14	AC101339	AC101339 Mus muscu	C 654	36	80.0	125378	8	AL391416	AL391416 Human DNA
C 582	36	80.0	56330	14	AL353694	AL353694 Continuation (4 of	C 655	36	80.0	125952	14	AC157881	AC157881 Monodelph
C 583	36	80.0	57313	5	BX784401	BX784401 Zebrafish	C 656	36	80.0	126306	14	AL391242	AL391242 Homo sapi
C 584	36	80.0	58250	14	AC103690	AC103690 Homo sapi	C 657	36	80.0	126356	14	AC006876	AC006876 Caenorhab
C 585	36	80.0	60917	14	AC090876	AC090876 Homo sapi	C 658	36	80.0	126750	8	AC010431	AC010431 Homo sapi
C 586	36	80.0	62841	14	AC100606	AC100606 Mus muscu	C 659	36	80.0	127462	15	F5D14	AC007767 Sequence
C 587	36	80.0	62841	14	AC100606	AC100606 Mus muscu	C 660	36	80.0	130296	8	HSJ3495K2	AL109924 Human DNA

c 661	36	80.0	130381	8	AC002402	AC002402 Human Chr	c 734	36	80.0	167511	14	AC083802	AC083802 Homo sapi
c 662	36	80.0	131398	8	H5445C9	Z95115 Human DNA	c 735	36	80.0	167900	14	AP001650	AP001650 Homo sapi
c 663	36	80.0	131501	8	AL139098	AL139098 Human DNA	c 736	36	80.0	168242	8	AC140830	AC140830 Homo sapi
c 664	36	80.0	131855	5	BX901910	BX901910 Zebrafish	c 737	36	80.0	168430	14	AC140830	AC140830 Homo sapi
c 665	36	80.0	134174	8	AC009495	AC009495 Homo sapi	c 738	36	80.0	168444	5	BX957335	BX957335 Zebrafish
c 666	36	80.0	136551	8	AC123786	AC123786 Homo sapi	c 739	36	80.0	169841	14	AC073888	AC073888 Homo sapi
c 667	36	80.0	136615	4	AC150730	AC150730 Atelari	c 740	36	80.0	170125	14	AC161298	AC161298 Oryzotag
c 668	36	80.0	137737	8	AC010378	AC010378 Homo sapi	c 741	36	80.0	170200	9	AC137743	AC137743 Mus muscu
c 669	36	80.0	138187	2	AC159438	AC159438 Trypanoso	c 742	36	80.0	170227	14	AC019161	AC019161 Homo sapi
c 670	36	80.0	138685	14	AC160234	AC160234 Rhinoloph	c 743	36	80.0	170233	14	AC068844	AC068844 Homo sapi
c 671	36	80.0	139490	15	AP003225	AP003225 Oryza sat	c 744	36	80.0	170279	14	AC132368	AC132368 Mus muscu
c 672	36	80.0	140245	15	AP005658	AP005658 Oryza sat	c 745	36	80.0	171041	14	AC148510	AC148510 Macropus
c 673	36	80.0	140403	14	AC109444	AC109444 Homo sapi	c 746	36	80.0	171645	5	BX957307	BX957307 Zebrafish
c 674	36	80.0	140978	2	AC099045	AC099045 Trypanoso	c 747	36	80.0	172266	2	AC092190	AC092190 Drosophi
c 675	36	80.0	141511	14	AC155475	AC155475 Zea mays	c 748	36	80.0	172321	8	AC109631	AC109631 Homo sapi
c 676	36	80.0	141620	14	AC142444	AC142444 Homo sapi	c 749	36	80.0	172543	8	AC106714	AC106714 Homo sapi
c 677	36	80.0	142175	14	CR759952	CR759952 Danio rer	c 750	36	80.0	172982	14	AC146434	AC146434 Pan trogl
c 678	36	80.0	144631	8	AC069027	AC069027 Homo sapi	c 751	36	80.0	173245	14	AL162380	AL162380 Homo sapi
c 679	36	80.0	145085	14	AC090261	AC090261 Homo sapi	c 752	36	80.0	173275	14	AC145031	AC145031 Homo sapi
c 680	36	80.0	145952	14	AC135921	AC135921 Oryza sat	c 753	36	80.0	173498	8	AC069575	AC069575 Homo sapi
c 681	36	80.0	147254	15	OSJN00150	AL662948 Oryza sat	c 754	36	80.0	173525	14	AC135022	AC135022 Rattus no
c 682	36	80.0	147352	15	AC137746	AC137746 Oryza sat	c 755	36	80.0	173738	8	AL136132	AL136132 Human DNA
c 683	36	80.0	147516	15	AP004622	AP004622 Oryza sat	c 756	36	80.0	174413	14	AC011788	AC011788 Homo sapi
c 684	36	80.0	148161	14	AC105742	AC105742 Felis cat	c 757	36	80.0	174669	9	AC127698	AC127698 Mus muscu
c 685	36	80.0	148320	14	AC026885	AC026885 Homo sapi	c 758	36	80.0	175090	9	AC122234	AC122234 Mus muscu
c 686	36	80.0	149562	14	AC143353	AC143353 Homo sapi	c 759	36	80.0	175651	8	AC145783	AC145783 Pan trogl
c 687	36	80.0	150150	15	AP002539	AP002539 Oryza sat	c 760	36	80.0	175681	2	AC007417	AC007417 Drosophi
c 688	36	80.0	150305	4	AC087423	AC087423 Sus scro	c 761	36	80.0	175960	9	AC151278	AC151278 Mus muscu
c 689	36	80.0	151332	9	AC122298	AC122298 Mus muscu	c 762	36	80.0	176074	9	AC100726	AC100726 Mus muscu
c 690	36	80.0	151717	9	AC162861	AC162861 Mus muscu	c 763	36	80.0	176145	14	AC027748	AC027748 Homo sapi
c 691	36	80.0	151857	8	AL831770	AL831770 Human DNA	c 764	36	80.0	176202	8	CNS05TF3	AL359951 Human chr
c 692	36	80.0	152024	8	AC063952	AC063952 Homo sapi	c 765	36	80.0	176204	8	AC119151	AC119151 Homo sapi
c 693	36	80.0	152134	5	BX571944	BX571944 Zebrafish	c 766	36	80.0	176395	2	AC007928	AC007928 Drosophi
c 694	36	80.0	152607	14	AC107321	AC107321 Felis cat	c 767	36	80.0	177446	8	AC147035	AC147035 Pan trogl
c 695	36	80.0	153368	15	AC136222	AC136222 Oryza sat	c 768	36	80.0	177576	2	AC022349	AC022349 Drosophi
c 696	36	80.0	153547	14	AL345327	AL345327 Homo sapi	c 769	36	80.0	177655	9	AC132288	AC132288 Mus muscu
c 697	36	80.0	153627	8	AC147054	AC147054 Pan trogl	c 770	36	80.0	177865	14	AC157455	AC157455 Monodelph
c 698	36	80.0	154142	9	AC132567	AC132567 Mus muscu	c 771	36	80.0	177882	8	AP000722	AP000722 Homo sapi
c 699	36	80.0	154384	14	AC015514	AC015514 Homo sapi	c 772	36	80.0	178184	9	AL683866	AL683866 Mouse DNA
c 700	36	80.0	154435	14	AC166781	AC166781 Mus muscu	c 773	36	80.0	178361	8	AC006042	AC006042 Homo sapi
c 701	36	80.0	154630	14	AC026059	AC026059 Homo sapi	c 774	36	80.0	179077	2	AC018489	AC018489 Drosophi
c 702	36	80.0	155332	14	AC040900	AC040900 Homo sapi	c 775	36	80.0	179384	9	AC122800	AC122800 Mus muscu
c 703	36	80.0	155519	9	AC118590	AC118590 Mus muscu	c 776	36	80.0	180169	14	CT010446	CT010446 Mus muscu
c 704	36	80.0	156700	14	AC027051	AC027051 Homo sapi	c 777	36	80.0	180496	14	CT0113875	AC113875 Rattus no
c 705	36	80.0	156829	8	AP003398	AP003398 Homo sapi	c 778	36	80.0	181236	14	AC132363	AC132363 Mus muscu
c 706	36	80.0	157165	4	AC087843	AC087843 Bos tauru	c 779	36	80.0	181236	14	AC022757	AC022757 Homo sapi
c 707	36	80.0	157171	14	AC155454	AC155454 Zea mays	c 780	36	80.0	181584	14	AC017027	AC017027 Homo sapi
c 708	36	80.0	157298	14	CR749747	CR749747 Danio rer	c 781	36	80.0	182303	8	AL356750	AL356750 Human DNA
c 709	36	80.0	158241	15	AP003409	AP003409 Oryza sat	c 782	36	80.0	183437	8	AC068933	AC068933 Homo sapi
c 710	36	80.0	159188	14	AC037432	AC037432 Homo sapi	c 783	36	80.0	183735	14	AC011080	AC011080 Homo sapi
c 711	36	80.0	159546	14	AC113419	AC113419 Homo sapi	c 784	36	80.0	183976	8	AC100821	AC100821 Homo sapi
c 712	36	80.0	159645	14	AC092026	AC092026 Homo sapi	c 785	36	80.0	184192	14	AP001154	AP001154 Homo sapi
c 713	36	80.0	160029	14	AC166354	AC166354 Mus muscu	c 786	36	80.0	185395	14	AC118699	AC118699 Mus muscu
c 714	36	80.0	160504	2	AC091207	AC091207 Drosophi	c 787	36	80.0	185430	14	AC025706	AC025706 Homo sapi
c 715	36	80.0	160892	9	AC122294	AC122294 Mus muscu	c 788	36	80.0	185603	8	AP001086	AP001086 Homo sapi
c 716	36	80.0	161100	14	AC025659	AC025659 Homo sapi	c 789	36	80.0	185818	9	AC158653	AC158653 Mus muscu
c 717	36	80.0	161759	9	AC115920	AC115920 Mus muscu	c 790	36	80.0	186165	8	AC112187	AC112187 Homo sapi
c 718	36	80.0	162341	14	AC163856	AC163856 Bos tauru	c 791	36	80.0	186306	14	AC134641	AC134641 Rattus no
c 719	36	80.0	162525	5	BX004981	BX004981 Zebrafish	c 792	36	80.0	186412	14	AC120972	AC120972 Zebrafish
c 720	36	80.0	162898	14	AC026953	AC026953 Homo sapi	c 793	36	80.0	186820	5	BX469921	BX469921 Zebrafish
c 721	36	80.0	162901	14	AC165184	AC165184 Muntiacus	c 794	36	80.0	186835	8	AC145378	AC145378 Pan trogl
c 722	36	80.0	163317	8	AC096564	AC096564 Homo sapi	c 795	36	80.0	186929	2	AC090657	AC090657 Mus muscu
c 723	36	80.0	164503	8	AC145859	AC145859 Pan trogl	c 796	36	80.0	186929	2	AC091201	AC091201 Drosophi
c 724	36	80.0	164801	14	AC144793	AC144793 Mus muscu	c 797	36	80.0	187378	8	AL357150	AL357150 Human DNA
c 725	36	80.0	164905	14	CR381592	CR381592 Danio rer	c 798	36	80.0	187637	14	AC046133	AC046133 Homo sapi
c 726	36	80.0	165052	9	AC101177	AC101177 Mus muscu	c 799	36	80.0	187947	8	CNS05TF6	AL354755 Human chr
c 727	36	80.0	165241	9	AC161455	AC161455 Mus muscu	c 800	36	80.0	188013	9	AC154351	AC154351 Mus muscu
c 728	36	80.0	165736	14	AC144352	AC144352 Homo sapi	c 801	36	80.0	188198	14	AC152842	AC152842 Muntiacus
c 729	36	80.0	166015	8	AC113194	AC113194 Homo sapi	c 802	36	80.0	188364	8	AC007784	AC007784 Homo sapi
c 730	36	80.0	166352	2	AC010069	AC010069 Drosophi	c 803	36	80.0	189054	4	AC135272	AC135272 Rattus no
c 731	36	80.0	166555	9	AC124366	AC124366 Mus muscu	c 804	36	80.0	189092	14	AC011759	AC011759 Drosophi
c 732	36	80.0	166761	14	AL159173	AL159173 Homo sapi	c 805	36	80.0	189462	2	AC155333	AC155333 Mus muscu
c 733	36	80.0	166772	8	AL589763	AL589763 Human DNA	c 806	36	80.0	189618	2	AC099029	AC099029 Drosophi

807	36	80.0	189705	9	AC131088	AC131088 Mus muscu	880	36	80.0	224342	14	AC112867	AC112867 Rattus no
808	36	80.0	189721	14	CR385047	Danio rer	881	36	80.0	224735	14	AC132661	AC132661 Rattus no
C 809	36	80.0	189792	14	AC152892	Ovis arie	C 882	36	80.0	225273	14	AC158882	AC158882 Bos tauru
C 810	36	80.0	189962	14	AC148165	AC148165 Zea mays	C 883	36	80.0	225381	14	AC120772	AC120772 Rattus no
C 811	36	80.0	190313	9	AC101816	Mus muscu	C 884	36	80.0	225738	9	AC102633	Mus muscu
C 812	36	80.0	190489	8	AC067794	AC067794 Homo sapi	C 885	36	80.0	226356	14	AC106919	AC106919 Rattus no
C 813	36	80.0	191096	8	AC008834	AC008834 Homo sapi	C 886	36	80.0	227876	14	CR933764	Danio rer
C 814	36	80.0	191659	8	AC069381	AC069381 Homo sapi	C 887	36	80.0	228009	14	AC128008	Rattus no
C 815	36	80.0	191671	9	AL672249	AL672249 Mouse DNA	C 888	36	80.0	228558	14	AC130156	Rattus no
C 816	36	80.0	192001	14	AC066837	AC066837 Homo sapi	C 889	36	80.0	228892	14	AC130156	Rattus no
C 817	36	80.0	193627	14	AC091427	AC091427 Mus muscu	C 890	36	80.0	228976	5	AC144489	AC144489 Gasterost
C 818	36	80.0	193638	14	AC164633	AC164633 Mus muscu	C 891	36	80.0	229057	14	AC098089	AC098089 Rattus no
C 819	36	80.0	193980	8	AL392087	AL392087 Human DNA	C 892	36	80.0	230213	9	AL845485	AL845485 Mouse DNA
C 820	36	80.0	194449	14	CR677617	CR677617 Danio rer	C 893	36	80.0	230765	14	CR847495	CR847495 Danio rer
C 821	36	80.0	194761	14	AC161541	AC161541 Bos tauru	C 894	36	80.0	232460	14	AC103188	AC103188 Rattus no
C 822	36	80.0	194871	14	AC155107	AC155107 Bos tauru	C 895	36	80.0	232741	2	AE003705	AE003705 Drosophill
C 823	36	80.0	195653	14	AC108127	AC108127 Homo sapi	C 896	36	80.0	233369	15	AC165171	AC165171 Zea mays
C 824	36	80.0	196976	9	AL845528	AL845528 Mouse DNA	C 897	36	80.0	233826	5	EX005336	EX005336 Zebrafish
C 825	36	80.0	197055	14	AC153319	AC153319 Ornithorh	C 898	36	80.0	233887	14	AC133687	AC133687 Rattus no
C 826	36	80.0	197471	8	AC142310	AC142310 Pan trogl	C 899	36	80.0	234114	14	AC111449	AC111449 Rattus no
C 827	36	80.0	197817	8	CNS057CS	AL355833 Human chr	C 900	36	80.0	235133	14	AC115375	AC115375 Rattus no
C 828	36	80.0	198578	9	AC134571	AC134571 Mus muscu	C 901	36	80.0	235887	14	AC106092	AC106092 Rattus no
C 829	36	80.0	198895	9	AC118025	AC118025 Mus muscu	C 902	36	80.0	236036	14	AC166174	AC166174 Mus muscu
C 830	36	80.0	198948	8	AC158399	AC158399 Mus muscu	C 903	36	80.0	236166	14	AC149754	AC149754 Bos tauru
C 831	36	80.0	199200	8	AL359633	AL359633 Human DNA	C 904	36	80.0	236263	14	AC098600	AC098600 Rattus no
C 832	36	80.0	200141	9	AL845305	AL845305 Mouse DNA	C 905	36	80.0	236809	14	AL356746	AL356746 Homo sapi
C 833	36	80.0	200556	5	EX004997	EX004997 Zebrafish	C 906	36	80.0	237010	14	CR855337	CR855337 Danio rer
C 834	36	80.0	200581	8	AC106794	AC106794 Homo sapi	C 907	36	80.0	237119	14	AC011740	AC011740 Drosophill
C 835	36	80.0	201070	14	AC026291	AC026291 Homo sapi	C 908	36	80.0	237684	14	AC111665	AC111665 Rattus no
C 836	36	80.0	201275	9	AC102613	AC102613 Mus muscu	C 909	36	80.0	238739	14	CR405708	CR405708 Danio rer
C 837	36	80.0	201376	14	AC012411	AC012411 Homo sapi	C 910	36	80.0	239308	14	AC107141	AC107141 Rattus no
C 838	36	80.0	201399	8	AC357149	AL357149 Human DNA	C 911	36	80.0	239363	14	AC121057	AC121057 Rattus no
C 839	36	80.0	201476	9	AL713372	AL713372 Mouse DNA	C 912	36	80.0	239405	14	AC153269	AC153269 Bos tauru
C 840	36	80.0	201513	14	AC156521	AC156521 Bos tauru	C 913	36	80.0	240209	14	AC130920	AC130920 Rattus no
C 841	36	80.0	201559	14	AC133788	AC133788 Rattus no	C 914	36	80.0	240481	2	AE003830	AE003830 Drosophill
C 842	36	80.0	202533	14	AC146464	AC146464 Salmiri b	C 915	36	80.0	240485	14	AC123014	AC123014 Rattus no
C 843	36	80.0	202729	14	AC163560	AC163560 Bos tauru	C 916	36	80.0	240676	14	AC120761	AC120761 Rattus no
C 844	36	80.0	203043	14	AC162564	AC162564 Bos tauru	C 917	36	80.0	240971	14	AC130253	AC130253 Rattus no
C 845	36	80.0	203246	14	AL356606	AL356606 Homo sapi	C 918	36	80.0	241050	14	AC119506	AC119506 Rattus no
C 846	36	80.0	203287	5	CR376756	CR376756 Zebrafish	C 919	36	80.0	241286	14	CR932077	CR932077 Danio rer
C 847	36	80.0	203926	9	AC153599	AC153599 Mus muscu	C 920	36	80.0	242687	14	AC159708	AC159708 Mus muscu
C 848	36	80.0	204302	14	CR405716	CR405716 Danio rer	C 921	36	80.0	242999	14	AC156907	AC156907 Bos tauru
C 849	36	80.0	204361	14	AP002443	AP002443 Homo sapi	C 922	36	80.0	243343	14	AC097337	AC097337 Rattus no
C 850	36	80.0	204557	14	AC098473	AC098473 Rattus no	C 923	36	80.0	243510	5	EX005415	EX005415 Zebrafish
C 851	36	80.0	205074	9	AC114423	AC114423 Mus muscu	C 924	36	80.0	243802	14	AC106520	AC106520 Rattus no
C 852	36	80.0	205430	14	AP002756	AP002756 Homo sapi	C 925	36	80.0	244042	14	AC103004	AC103004 Rattus no
C 853	36	80.0	205436	9	AL671916	AL671916 Mouse DNA	C 926	36	80.0	244250	14	AC095307	AC095307 Rattus no
C 854	36	80.0	205587	5	EX294179	EX294179 Zebrafish	C 927	36	80.0	244573	14	AC133400	AC133400 Rattus no
C 855	36	80.0	207155	9	AC132431	AC132431 Mus muscu	C 928	36	80.0	244678	14	AC128751	AC128751 Rattus no
C 856	36	80.0	208029	9	AC105301	AC105301 Mus muscu	C 929	36	80.0	245795	14	AC096419	AC096419 Rattus no
C 857	36	80.0	208059	14	AC114192	AC114192 Rattus no	C 930	36	80.0	246456	14	AC095509	AC095509 Rattus no
C 858	36	80.0	208393	14	AC087870	AC087870 Mus muscu	C 931	36	80.0	247964	14	AC137320	AC137320 Rattus no
C 859	36	80.0	210079	14	AC160423	AC160423 Bos tauru	C 932	36	80.0	248019	14	AC109007	AC109007 Rattus no
C 860	36	80.0	210950	8	AC157954	AC157954 Pan trogl	C 933	36	80.0	248243	14	AC094831	AC094831 Rattus no
C 861	36	80.0	211061	9	AC153591	AC153591 Mus muscu	C 934	36	80.0	248419	5	EX470164	EX470164 Zebrafish
C 862	36	80.0	212364	9	AL691472	AL691472 Mouse DNA	C 935	36	80.0	248528	14	AC116084	AC116084 Rattus no
C 863	36	80.0	212617	9	AC117681	AC117681 Mus muscu	C 936	36	80.0	248983	9	AC124777	AC124777 Mus muscu
C 864	36	80.0	213178	14	AC164271	AC164271 Bos tauru	C 937	36	80.0	250022	14	AC107154	AC107154 Rattus no
C 865	36	80.0	214406	14	AC140672	AC140672 Mus muscu	C 938	36	80.0	250029	8	AE014309	AE014309 Homo sapi
C 866	36	80.0	214567	9	AC105305	AC105305 Mus muscu	C 939	36	80.0	250375	14	AC131167	AC131167 Rattus no
C 867	36	80.0	215899	14	CR932998	CR932998 Danio rer	C 940	36	80.0	250636	14	AC109056	AC109056 Rattus no
C 868	36	80.0	216334	8	AC011132	AC011132 Homo sapi	C 941	36	80.0	251179	14	AC139890	AC139890 Rattus no
C 869	36	80.0	216506	14	AC107127	AC107127 Rattus no	C 942	36	80.0	251455	14	AC111737	AC111737 Rattus no
C 870	36	80.0	216816	14	CR759846	CR759846 Danio rer	C 943	36	80.0	251457	2	AE003566	AE003566 Drosophill
C 871	36	80.0	216932	9	AC090656	AC090656 Mus muscu	C 944	36	80.0	252445	14	AC108348	AC108348 Rattus no
C 872	36	80.0	217109	14	AC158549	AC158549 Mus muscu	C 945	36	80.0	252789	14	AC095119	AC095119 Rattus no
C 873	36	80.0	217109	14	AC158549	AC158549 Mus muscu	C 946	36	80.0	254035	14	AC122974	AC122974 Rattus no
C 874	36	80.0	218657	14	AC156693	AC156693 Bos tauru	C 947	36	80.0	254794	14	CR925716	CR925716 Danio rer
C 875	36	80.0	219076	9	AL731713	AL731713 Mouse DNA	C 948	36	80.0	254959	14	AC123016	AC123016 Rattus no
C 876	36	80.0	220807	14	AC139643	AC139643 Rattus no	C 949	36	80.0	255446	14	AC099127	AC099127 Rattus no
C 877	36	80.0	222685	14	AC140971	AC140971 Canis fam	C 950	36	80.0	255663	9	AC091276	AC091276 Mus muscu
C 878	36	80.0	223068	9	AC123057	AC123057 Mus muscu	C 951	36	80.0	256073	8	AE006464	AE006464 Homo sapi
C 879	36	80.0	223830	9	AC153862	AC153862 Mus muscu	C 952	36	80.0	256110	14	AC116951	AC116951 Mus muscu


```

c 953 36 80.0 256675 5 BX511225
c 954 36 80.0 260560 14 AC137180
c 955 36 80.0 262545 14 AC126287
c 956 36 80.0 265793 14 AC115477
c 957 36 80.0 269101 14 AC098339
c 958 36 80.0 273468 2 AB003574
c 959 36 80.0 276888 14 AC151080
c 960 36 80.0 283826 14 AC113658
c 961 36 80.0 288446 14 AC161903
c 962 36 80.0 298906 14 AC157302
c 963 36 80.0 300468 2 AB003503
c 964 36 80.0 303299 14 AC119117
c 965 36 80.0 305600 1 CNSPAX05
c 966 36 80.0 315463 14 AC122070
c 967 36 80.0 337658 9 AC121114
c 968 36 80.0 338760 14 AC123485
c 969 36 80.0 339193 14 AC098394
c 970 36 80.0 349907 1 BX571874
c 971 36 80.0 349980 6 AX041921
c 972 35 77.8 148 10 G44076
c 973 35 77.8 250 10 G14730
c 974 35 77.8 274 10 BV296819
c 975 35 77.8 362 15 HFU296484
c 976 35 77.8 362 15 HPZ296485
c 977 35 77.8 364 15 AJ748787
c 978 35 77.8 415 6 AX410480
c 979 35 77.8 442 2 TCAS44858
c 980 35 77.8 443 2 ARE544853
c 981 35 77.8 445 2 AY059200
c 982 35 77.8 466 6 CQ098123
c 983 35 77.8 466 6 CQ136997
c 984 35 77.8 466 6 CQ220319
c 985 35 77.8 466 6 CQ258668
c 986 35 77.8 466 6 CQ296132
c 987 35 77.8 466 6 CQ332648
c 988 35 77.8 473 6 AR205037
c 989 35 77.8 519 10 BV298895
c 990 35 77.8 552 10 BV296508
c 991 35 77.8 574 10 BV325647
c 992 35 77.8 581 10 BV251051
c 993 35 77.8 593 10 BV021891
c 994 35 77.8 601 6 AR663750
c 995 35 77.8 602 10 BV276246
c 996 35 77.8 630 10 BV399270
c 997 35 77.8 632 2 AY227106
c 998 35 77.8 639 10 BV287069
c 999 35 77.8 656 10 BV510218
c1000 35 77.8 700 6 AR644631

```

ALIGNMENTS

```

RESULT 1
CQ687716
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

```

Sequence 32642 from Patent WO02070737. linear PAT 03-FEB-2004

CQ687716

CQ687716.1 GI:42218962

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

Liew, C.C., Marshall, W.B. and Zhang, H.

Compositions and methods relating to osteoarthritis

Patent: WO 02070737-A 32642 12-SEP-2002;

Chondrogene Inc. (CA)

Location/Qualifiers

1. .290

/organism="Homo sapiens"

/mol_type="unassigned DNA"

```

ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

```

2.02 45.00 290

100.0% 0

100.0% 0

100.0% 0

6 0

US-10-774-176-14 (1-9) x CQ687716 (1-290)

QY 1 ValLeuTyrLeuAenArgLysGlyIle 9

DB 98 GTTTTGTATTTCACCGCAAGGGGATA 124

RESULT 2

HSPA32B9/c HSPA32B9 421 bp DNA linear STS 21-MAY-1998

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Location/Qualifiers

1. .421

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="6"

/clone="SC6PA32B9"

/sex="female"

/tissue type="EBV lymphoblastoid cell line"

/clone_lib="SC6PA"

/dev_stage="adult"

/note="The estimated purity of the flow-sorted chromosome 6 library is >97%"

US-10-774-176-14 (1-9) x HSPA32B9 (1-421)

QY 1 ValLeuTyrLeuAenArgLysGlyIle 9

DB 196 GTTTTGTATTTCACCGCAAGGGGATA 170

RESULT 3

CQ920916

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

Sequence 2116 from Patent WO2004097052.

CQ920916

CQ920916.1 GI:56210857

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

Sequence 2116 from Patent WO2004097052.

CQ920916

CQ920916.1 GI:56210857

FEATURES	source	FT	Key	Location/Qualifiers	FT	Key	Location/Qualifiers
ORIGIN							
Alignment Scores:							
Pred. No.:			6.31	Length:		901	
Score:			45.00	Matches:		9	
Percent Similarity:			100.0%	Conservative:		0	
Best Local Similarity:			100.0%	Mismatches:		0	
Query Match:			100.0%	Indels:		0	
DB:			6	Gaps:		0	
US-10-774-176-14 (1-9) x BD249733 (1-901)							
Qy	1	ValleuTyrlEuAsnArgLySGlyIle	9				
Db	684	GTTTTGATTGTGAACCGCAAGGGGATA	710				
RESULT 5							
AX025013							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
Canis sp.							
Canis sp.							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;							
Canis.							
REFERENCE							
1							
AUTHORS							
TITLE							
JOURNAL							
Carroll, M.W. and Myers, K.A.							
Polypeptide							
Patent: WO 0029428-A 3 25-MAY-2000;							
CARROLL MILLS WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD							
BIOMEDICA LTD (GB)							
FEATURES							
source							
1.							
Location/Qualifiers							
/organism="Canis sp."							
/mol_type="unassigned DNA"							
/db_xref="taxon:9616"							
ORIGIN							
Alignment Scores:							
Pred. No.:			6.31	Length:		901	
Score:			45.00	Matches:		9	
Percent Similarity:			100.0%	Conservative:		0	
Best Local Similarity:			100.0%	Mismatches:		0	
Query Match:			100.0%	Indels:		0	
DB:			6	Gaps:		0	
US-10-774-176-14 (1-9) x AX025013 (1-901)							
Qy	1	ValleuTyrlEuAsnArgLySGlyIle	9				
Db	684	GTTTTGATTGTGAACCGCAAGGGGATA	710				
RESULT 6							
AX316088							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
Canis sp.							
Canis sp.							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Canis.							
REFERENCE							
1							
AUTHORS							
TITLE							
JOURNAL							
Carroll, M.W. and Myers, K.A.							
Polypeptide				</			

Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 3 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
1..901
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN

Alignment Scores:
Pred. No.: 6.31 Length: 901
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX316088 (1-901)

Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9
|||||
Db 684 GTTTTGATTGACCGCAAGGGGATA 710

RESULT 7
AX829164
LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1
AUTHORS Mack, D.H., Gish, K.C. and Afar, D.
TITLE Methods of diagnosis of breast cancer, compositions and methods of screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)

FEATURES
source
1..927
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 6.49 Length: 927
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX829164 (1-927)

Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9
|||||
Db 775 GTTTTGATTGACCGCAAGGGGATA 801

RESULT 8
AX467373
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603

KEYWORDS
SOURCE Felis sp.
ORGANISM Felis sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.

REFERENCE 1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
1..1260
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"

ORIGIN

Alignment Scores:
Pred. No.: 8.83 Length: 1260
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX467373 (1-1260)

Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9
|||||
Db 1114 GTTTTGATTGACCGCAAGGGGATA 1140

RESULT 9
AX821533
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929

KEYWORDS Felis catus (cat)
SOURCE Felis catus
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.

REFERENCE 1
AUTHORS Carroll, M.M., Kingman, S.M. and Redchenko, I.M.
TITLE MHC class I peptide epitopes from the human 5t4 tumor-associated antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN

Alignment Scores:
Pred. No.: 8.83 Length: 1260
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX821533 (1-1260)

Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9
|||||
Db 1114 GTTTTGATTGACCGCAAGGGGATA 1140

RESULT 10
AX821548

LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS Felis catus (cat)
SOURCE Felis catus
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE 1
AUTHORS Carroll, M.O., Harrop, R.O. and Kingsman, S.O.
TITLE MHC class II peptide epitope of 54 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
Location/Qualifiers
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 8.83 Length: 1260
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x AX821548 (1-1260)
QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
|||||
Db 1114 GTTTGTACTTGACCGCAAGGGGATA 1140
RESULT 11
BD249731
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1263)
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K7/06, C07K14/065,
C07K19/00,
PC C12N15/00
CC Polypeptide
FH Key
FT source Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 8.86 Length: 1263
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x BD249731 (1-1263)
QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
|||||
Db 1117 GTTTGTATTGACCGCAAGGGGATA 1143
RESULT 12
AX025011 1263 bp DNA linear PAT 15-SEP-2000
LOCUS AX025011
DEFINITION Sequence 1 from Patent WO029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL Patent: WO 029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB); MYERS KEVIN ALAN (GB); OXFORD
BIOMEDICA LTD (GB)
FEATURES
Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 8.86 Length: 1263
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x AX025011 (1-1263)
QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
|||||
Db 1117 GTTTGTATTGACCGCAAGGGGATA 1143
RESULT 13
AX149553
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kingsman, A.O., Kingsman, S.M., Bebbington, C.R., Carroll, M.W.,
Ellard, F.M. and Myers, K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
Location/Qualifiers
1..1263
/organism="synthetic construct"

/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="5T4"

ORIGIN

Alignment Scores:
Pred. No.: 8.86 Length: 1263
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX149553 (1-1263)

Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9

Db 1117 GTTTTGATTGACCGCAAGGGGATA 1143

RESULT 14

AX316086 AX316086 1263 bp DNA linear PAT 14-DEC-2001
LOCUS Sequence 1 from Patent EP1160323.
DEFINITION AX316086
ACCESSION AX316086
VERSION AX316086.1 GI:17899278

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 Carroll, M.W. and Myers, K.A.
St4 tumour-associated antigen for use in tumour immunotherapy

TITLE Patent: EP 1160323-A 1 05-DEC-2001;

JOURNAL Oxford Biomedica (UK) Limited (GB)

FEATURES

source
1..1263
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 8.86 Length: 1263
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX316086 (1-1263)

Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9

Db 1117 GTTTTGATTGACCGCAAGGGGATA 1143

RESULT 15

AX467371 AX467371 1263 bp DNA linear PAT 16-JUL-2002
LOCUS Sequence 1 from Patent WO0238612.
DEFINITION AX467371
ACCESSION AX467371

VERSION

AX467371.1 GI:21900602

KEYWORDS

SOURCE Canis sp.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE

1 Myers, K., Drury, N. and Carroll, M.
Polypeptide

TITLE Patent: WO 0238612-A 1 16-MAY-2002;

JOURNAL

1135 GTTTTGATTGACCGCAAGGGGATA 1161

FEATURES
source
1..1263
Location/Qualifiers
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN

Alignment Scores:
Pred. No.: 8.86 Length: 1263
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX467371 (1-1263)

Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9

Db 1117 GTTTTGATTGACCGCAAGGGGATA 1143

RESULT 16

BD249732 BD249732 1281 bp DNA linear PAT 17-JUL-2003
LOCUS Polypeptide.
DEFINITION BD249732
ACCESSION BD249732
VERSION BD249732.1 GI:33059502

KEYWORDS

JP 2002530060-A/2.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 1281)

AUTHORS

Carroll, M.W. and Myers, K.A.

TITLE

Polypeptide

JOURNAL

Patent: JP 2002530060-A 2 17-SEP-2002;

COMMENT

OXFORD BIOMEDICA LTD

OS Mus musculus (mouse)

PN JP 2002530060-A/2

PD 17-SEP-2002

PF 18-NOV-1999 JP 2000582415

PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR

30-JUL-1999 GB 9917995.4

PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS

PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K14/065,

PC C07K19/00,

PC C12N15/00

CC Polypeptide

FH Key

FT source

1..1281

Location/Qualifiers

/organism="Mus musculus (mouse)"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

FEATURES

source

1..1281

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

ORIGIN

Alignment Scores:

Pred. No.: 8.98 Length: 1281
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x BD249732 (1-1281)

Qy 1 ValLeuTyrLeuAsnArgLysGlyIle 9

Db 1135 GTTTTGATTGACCGCAAGGGGATA 1161

```

RESULT 17
AX025012
LOCUS AX025012 1281 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 2 from Patent WO0029428.
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 2 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source
1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.: 8.98 Length: 1281
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x AX025012 (1-1281)

Qy 1 ValLeuTyrrLeuAaAnArgLySGLyIle 9
Db 1135 GTTTTGATTGTAACCGTAAGGCATA 1161

RESULT 18
AX316087
LOCUS AX316087 1281 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 2 from Patent EP1160323.
ACCESSION AX316087
VERSION AX316087.1 GI:17899279
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 2 05-DEC-2001;
OXFORD Biomedica (UK) Limited (GB)
FEATURES
source
1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.: 8.98 Length: 1281
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-14 (1-9) x AX316087 (1-1281)

```

```

Qy 1 ValLeuTyrrLeuAaAnArgLySGLyIle 9
Db 1135 GTTTTGATTGTAACCGTAAGGCATA 1161

RESULT 19
HS5T4OA
LOCUS HS5T4OA 2053 bp RNA linear PRI 18-APR-2005
DEFINITION Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.
ACCESSION Z29083
VERSION Z29083.1 GI:435654
KEYWORDS 5T4 gene; 5T4 oncofoetal antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Myers,K.A., Rahi-Saund,V., Davison,M.D., Young,J.A., Cheater,A.J.
and Stern,P.L.
TITLE Isolation of a cDNA encoding 5T4 oncofoetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
JOURNAL J. Biol. Chem. 269 (12), 9319-9324 (1994)
PUBMED 8132670
REFERENCE 2 (bases 1 to 2053)
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
FEATURES
source
1..2053
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label=N-flank
85..1347
/codon_start=1
/evidence=experimental
/product="5T4 oncofoetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGSRGPAAGDGRLLRLARLALVLLGWVSSSPSSPTSSASFSSS
APFLASAVSAQPLPDQCPCSEARTVKVNRNLTEVDTDLPAYVYRNLFTGNQ
LAVLPAGAPAPPLAEALNLSGRLDEVRAGAPEHLPSLRQLDLSHNLADLSPP
AFSGSNASVAPSPLVELTNHVPEDERONRSPEGVMVAALLAGRALQGLRLRLIA
SNHFLYLPDVLQPLSLRHLDLNNLSVLTYSVSPRNLTHLSLEHLDNALKVLNG
TLAEQGLPHIRVFLDNPNVCDCHMADMTWLKSTEVQGGKRLTCAYPEKRRNRL
LELNSADLDCDPLPPSQTSYVFLGIVLALIGAIFLLVLYLNKRGKIKKMMHNRDAC
RDHMEGYHYRYBINADPRLTNLSSNSDV"
130..171
sig_peptide
misc_RNA 373..966
/product="Leucine rich repeat region"
/label=LRRS
966..1119
/misc_RNA /product="LRR C-terminal flank"
/label=C-flank
1153..1215
/misc_RNA /product="transmembrane peptide"
/standard_names="transmembrane region"
/function="Anchorage of the protein to the cell membrane"
ORIGIN

```

Alignment Scores:

Pred. No.: 14.4 Length: 2053
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x HS5740A (1-2053)

Qy 1 ValLeuTyfLeuAenArgLyseGlyle 9
 |||||
 Db 1201 GTTTTGTATTGAACCGAAGGGGATA 1227

RESULT 20

LOCUS CR855786 2183 bp mRNA linear VRT 03-NOV-2004
 DEFINITION Xenopus tropicalis finished cDNA, clone TGas020h08.
 ACCESSION CR855786

VERSION CR855786.1 GI:55295318

KEYWORDS Xenopus tropicalis (Silurana tropicalis)

SOURCE Xenopus tropicalis
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS 1 (bases 1 to 2183)
 Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
 Francis,M.D., Garrett,N., Gilchrist,M.J., Graham,D.V.,
 McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
 Voigt,J. and Zorn,A.M.

TITLE

JOURNAL Submitted (03-NOV-2004) Sanger Institute, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: tropesanger.ac.uk

COMMENT

This sequence is from a Xenopus Gene Collection (XGC) library, from
 a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
 extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
 ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
 end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI

Host: Escherichia coli XL1-blue.

FEATURES

source
 Location/Qualifiers

1..2183
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TGas020h08"
 /clone_lib="XGC-gastrula"
 /dev_stage="gastrula (stage 10.5-13 mixed)"

ORIGIN

Alignment Scores:
 Pred. No.: 15.3 Length: 2183
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0

US-10-774-176-14 (1-9) x CR855786 (1-2183)

Qy 1 ValLeuTyfLeuAenArgLyseGlyle 9
 |||||
 Db 792 GTTTTATATTAAACAGGAAGGGATT 818

RESULT 21

AF063939
 LOCUS Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA, ROD 01-JAN-2000
 DEFINITION complete cds.

ACCESSION

AF063939

VERSION AF063939.1 GI:6650211

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 2333)

Ninkina,N.N. and Buchman,V.L.

Structure and expression of the rat 5T4 gene

TITLE

JOURNAL Unpublished

2 (bases 1 to 2333)

Buchman,V.L.

Direct Submission

Submitted (06-MAY-1998) School of Biomedical Sciences, University

of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,

UK

FEATURES

source

Location/Qualifiers

1..2333
 /organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/tissue_type="cerebellum"

/dev_stage="newborn"

1..2333

/gene="5T4"

1..363

/gene="5T4"

364..1644

/gene="5T4"

/codon_start=1

/product="5T4 oncofetal antigen homolog"

/protein_id="AAF21770.1"

/db_xref="GI:6650212"

/translation="MPGAGSGPSAGDGLRLARLALVLLGWVSAPSSSLPSSSTS

PAAFIAGSAGPPPAERCPACSEAAATKCVNRNLLEVPADLPVYVNLFLTNQ

MTVLPAAGAPQPPADLAVLNLGSHKEVGAGAFELHGLRLDLNLSNPITNLGAF

TPAGSNVSTPSPLELILNIHVPPQQRQSGFEGMVAFSGMAAALRSGGLRGL

HLLEASHNHYLYPRDLIDQLPSLHLDLRNNSLVSTYASFRNLTHLSLHLENDAL

KVHNSLTAEWQGLAHVRVFLDNNPWCDVMADVMVSLKETETVPDKARLTCAFPFK

MNRGLDLTSSDLDCDATTLOSLOTSSVFLGIVLALIGAILFLVLYLNKRGKKMWH

NIRDCRDMEGHYRYEINADPSLTNLSNSDV"

1645..2333

/gene="5T4"

2315..2320

/gene="5T4"

Alignment Scores:

Pred. No.: 16.4 Length: 2333

Score: 45.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 9 Gaps: 0

US-10-774-176-14 (1-9) x AF063939 (1-2333)

Qy 1 ValLeuTyfLeuAenArgLyseGlyle 9

|||||

Db 1498 GTTTTGTATTGAACCGAAGGGGATA 1524

RESULT 22

BD127282

LOCUS Primer for synthesizing full-length cDNA and use thereof.

DEFINITION BD127282

ACCESSION BD127282

VERSION BD127282.1 GI:23222227

KEYWORDS JP 2002017375-A/2713.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayaashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2713 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2713
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
(424)...(1572).
source
1..2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 16.6 Length: 2359
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x BD127282 (1-2359)

Qy 1 ValLeuTyrrLeuAsnArgLysGlyIle 9
|||||
Db 1540 GTTTTGTATTGAACCGCAAGGGGATA 1566

RESULT 23
CQ782724
LOCUS Sequence 2864 from Patent EP1396543. linear PAT 17-MAR-2004
DEFINITION CQ782724
ACCESSION CQ782724.1 GI:45502667
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
Ota,T., Nishikawa,T., Isogai,T., Hayaashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2864 10-MAR-2004; (JP)
RESEARCH Association for Biotechnology (JP)
FEATURES
Location/Qualifiers
1..2359
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAP85958.1"

CDS

```

```

/db_xref="GI:45502668"
/translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSSPTSSASSPSS
APFLSAVSAQFPPLPQCPALCESSAARTVKCVNRLTEVPTDLPAYVRLFTGQ
LAVLPAGAFARPPPLAELALNLSRLDEVRAGAFEHLPSTRLDLSHNPDLADLSPF
AFSGSNASVSAFSPVLVELILNHITVPPEQRQNRSPFGMVVAALLAGRALQGLRLELA
SNHFLYLPDVLQALPSLRHLDLNNLSVSLTVVSPFNLTHLSLEHEDNALKVLHNG
TLAEQGLPHIRVFLDNNPVCDCMADMTWLKETEVVGKDRLTCAVPEKMRNRL
LELSADLDCDFILPFSLQTSVYFGLVIALGAIFFLLVLYLNKGIKK"

ORIGIN
Alignment Scores:
Pred. No.: 16.6 Length: 2359
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x CQ782724 (1-2359)

Qy 1 ValLeuTyrrLeuAsnArgLysGlyIle 9
|||||
Db 1540 GTTTTGTATTGAACCGCAAGGGGATA 1566

RESULT 24
AK074786
LOCUS Homo sapiens cDNA FLJ90305 fis, clone NT3RP2000694, highly similar
to Homo sapiens 574 oncofetal trophoblast glycoprotein gene.
DEFINITION AK074786 2359 bp mRNA linear PRI 03-SEP-2002
ACCESSION AK074786
VERSION AK074786.1 GI:22760460
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
Isogai,T., Ota,T., Nishikawa,T., Hayaashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 2359)
Isogai,T. and Otsuki,T.
Direct Submission
SUBMITTED (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
Location/Qualifiers
1..2359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT3RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT3RP2"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

```

Alignment Scores:
 Pred. No.: 16.6 Length: 2359
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x AK074786 (1-2359)

Qy 1 ValLeuTyTLLeuAsnArgLysGlyTle 9
 Db 1540 GTTTGTATTGACCGCAAGGGGATA 1566

RESULT 25

BD127283

LOCUS BD127283 2361 bp DNA linear PAT 18-SEP-2002

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD127283

VERSION BD127283.1 GI:23222228

KEYWORDS JP 2002017375-A/2714.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Iehii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.

Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002017375-A 2714 22-JAN-2002;

HEILX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002017375-A/2714

PD 22-JAN-2002

PF 07-JUL-2000 JP 2000253172

PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO

PI ISHII,

PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI

SHINICHI KOJIMA,

PI TETSUUI OTSUKI, HISASHI KOGA

PC

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers (426). (1685).

FT CDS

Location/Qualifiers

1..2361

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 16.6 Length: 2361
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x BD127283 (1-2361)

Qy 1 ValLeuTyTLLeuAsnArgLysGlyTle 9

Db 1542 GTTTGTATTGACCGCAAGGGGATA 1568

RESULT 26

CQ782726

LOCUS

CQ782726 2361 bp DNA linear PAT 17-MAR-2004

DEFINITION Sequence 2866 from Patent EP1396543.

ACCESSION CQ782726

VERSION CQ782726.1 GI:45502669

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Iehii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.

Primer for synthesizing full length cDNA clones and their use
 Patent: EP 1396543-A 2866 10-MAR-2004;

Research Association for Biotechnology (JP)

FEATURES

source

1..2361

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

426..1688

/note="unnamed protein product"

/codon_start=1

/protein_id="CAF85961.1"

/translation="MPGCGSRGPAAGDGRRLRLARLALLVLLGWSSSPTSSASSFSSSS

APFLASVSAQPLPDQPCALCECEAAATVKNRNLTEVPTDLPAYVRLPLTGNQ

LAVLPAGAFARPPPLAEALNLSGRSLDEVRAGAEHLPSLRQLDLSHPLADLSPP

AFSGSNASVSAPSLVELILNHIIVPEPERQNRSEGMVVAALLAGRLRLLELA

SNHFLYLPRLVLAQPSLRHLDSNLSVSLTVFSRGNLTHLESJHLEDNALKVLHNG

TLAEIQGLPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPSKMRNRL

LELNSADLDCDPLPSPSQTSYVFLGIVLALIGALFLVLVLRNRGIKKMMNIRDAC

RDMESGYHYRYEINADPRLTNLSSSDV"

ORIGIN

Alignment Scores:

Pred. No.: 16.6 Length: 2361
 Score: 45.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x CQ782726 (1-2361)

Qy 1 ValLeuTyTLLeuAsnArgLysGlyTle 9

Db 1542 GTTTGTATTGACCGCAAGGGGATA 1568

RESULT 27

AX961916

LOCUS

AX961916

DEFINITION

Sequence 127 from Patent WO03104277.

ACCESSION

AX961916

VERSION

AX961916.1 GI:40881326

KEYWORDS

source

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.

TITLE

Stat6 activation gene

JOURNAL

Patent: WO 03104277-A 127 18-DEC-2003;

Asahi Kasei Kabushiki Kaisha (JP)

FEATURES

source

1..2361

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

426..1688

CDS

/note="unnamed protein product"

/codon_start=1

/protein_id="CAF06467.1"

/db_xref="GI:40881327"

/translation="MFGGSGRPAAGDGRRLRLARLALVLLGVSSPTSSSSPSSS
APFLASVAGPPLDQPCALCESEAATKCVNRNLTEVPTDLPAYVYRNFLPTGNQ
LAVLPAGAFARPPLEALAAALNSSRLDEVRAGAFELPLSLRLDLSHPLADLSPF
AFSGSNASVSPPLVELLNHIVPDERQRNSRFGMVVAALLAGRALQGLRLLELA
SNHFLYLPRLVLAQPSLRHLDSNLSVLSVTSVSPRNITLHSLHLEDAKLVHLNG
TLAEIQGLPHRVLFDNPNFVCDCHMADMTWLKETEYVQGRDLTCAYPEKRRNVL
LELNSADLDCDPLPSPQTSYVFLGIVLALIGAIFLLVLYLNKGIKKWMENIRDAC
RDHMEGYHYRYEINADPRLTNLSSSDV"

ORIGIN

Alignment Scores:
Pred. No.: 16.6 Length: 2361
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX961916 (1-2361)

QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
DB 1542 GTTTTGTATTGTAACCGCAAGGGGATA 1568

RESULT 28

AK074790

LOCUS

DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar to Homo sapiens 574 oncofetal trophoblast glycoprotein gene.

ACCESSION AK074790.1 GI:22760466

VERSION AK074790.1

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Otsuki,T., Ota,T., Nishikawa,T., Hayaashi,K., Suzuki,Y., Yamamoto,J., Wakamatsu,A., Kimura,K., Sakamoto,K., Hatanano,N., Kawai,Y., Ishii,S., Saito,K., Kojima,S., Sugiyama,T., Ono,T., Okano,K., Yoshikawa,Y., Aotaka,S., Sasaki,N., Hattori,A., Okumura,K., Nagai,K., Sugano,S., and Isogai,T.
Signal Sequence and Keyword Trap in silico for Selection of Full-length Human cDNAs Encoding Secretion or Membrane Proteins from Oligo-Capped cDNA Libraries
DNA Res. 12, 117-126 (2005)

JOURNAL

REFERENCE

AUTHORS

Isogai,T., Ota,T., Nishikawa,T., Hayaashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotaka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Ninomiya,K.
NEDO human cDNA sequencing project
Unpublished

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).

FEATURES

source

Location/Qualifiers

1..2361
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/notes="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 16.6 Length: 2361
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x AK074790 (1-2361)

QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9
DB 1542 GTTTTGTATTGTAACCGCAAGGGGATA 1568

RESULT 29

BC087011

LOCUS

DEFINITION Rattus norvegicus trophoblast glycoprotein, mRNA (cDNA clone MGC:93332 IMAGE:7193411), complete cds.

ACCESSION BC087011

VERSION BC087011.1 GI:56268819

KEYWORDS MGC.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Murinae; Rattus.

REFERENCE

AUTHORS

Krausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Strausner,R.D., Collins,F.S., Wagner,L., Shermen,C.M., Schuler,G.D., Altschul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,P., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McGowan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Sanchez,J., Helton,E., Kettman,M., Madan,A.C., Rodrigues,S., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

Email: cgapbs-remail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 186 Row: 0 Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES
source
1..2361
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93332 IMAGE:7193411"
/tissue_type="Heart, rat (Brown Norway)"
/clone_lib="NIH_MGC_234"
/lab_host="DH10B"
/note="Vector: pExpress1"

gene
1..2361
/gene="Tpbg"
/note="synonym: 574"
/db_xref="GeneID:83684"
/db_xref="RGD:621453"
364..1644
/gene="Tpbg"
/codon_start=1
/product="Tpbg protein"
/protein_id="AAH87011.1"
/db_xref="GI:56268820"
/db_xref="GeneID:83684"
/db_xref="RGD:621453"
/tranlation="MPGAGSRGPGAGDGLRLRLALVLLGWVSASAPSSISPSSTS
PAAFLAGSAGPPAPCPACSEAAATKVCNVRNLEVPADLPVVRNIFLGNQ
MTVLPAAGAFAPPLADLANLNGHLKEVGAGAFHLPLGLRLDLSHNPLNLISAF
TFAGNSVSTPSPLEILHIVPPEDQRQNGSPGVAFEGMVAAALRGLALRGL
HHLBSAHFYLPRDLIDQLPSLKHLDLRNNSLVSTYASFRNLTHLESLELDNAL
KVLHNSGLAEQGLAHVRFLDNNPWCDCYMAWMVSWLKSTEVVPDKARLTCAPEK
MRNGLDLTSSDLDCDRTLQSLQTSVFGLGIVLALGALFLVLYLNRGKIKKWMH
NIRDACRDHMGTHRYEINADPRUTNLSSSDV"

CDS
16.6 Length: 2361
45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

ORIGIN
US-10-774-176-14 (1-9) x BC087011 (1-2361)
Oy 1 ValLeuTyrLeuAsnArgLysGlyIle 9
|||||
Db 1498 GTTTGTATTGTAACCGGAAGGCATA 1524
|||||
RESULT 30
BC037161
LOCUS
DEFINITION BC037161 2379 bp mRNA linear PRI 29-JUN-2004
Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317
IMAGE:4138906), complete cds.
ACCESSION BC037161
VERSION BC037161.2 GI:33872201
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2379)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Sheets, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richard, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Pahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2379)
Strausberg, R.
Direct Submission
Submitted (03-SEP-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:22713382.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 26 Row: m Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5729717.

FEATURES
source
1..2379
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:15317 IMAGE:4138906"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2379
/gene="TPBG"
/note="synonyms: M6P1, 5T4-AG, 5T4"
/db_xref="GeneID:7162"

/db_xref="MIM:190920"
427..1589
/gene="tpbg"
/codon_start=1
/product="574 oncofetal trophoblast glycoprotein"
/protein_id="AAH37161.1"
/db_xref="GI:22713383"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"
/translation="MPGCGRGPAAAGDRLRLARLALVLLGVWSSSPSSPTSSSPSS
APFLASVAGPPLPQPCALCESEARTKVCNRLNLEVPADLPYVYRNLPLTGNQ
LAVLPAGAPARQPLAEALAAALNSGRLSDVDEVRAGAFELPSRLQDLNPLADLSPF
AFSGSNASVSPLEELLNHIVPPEDORQNGSPGMAFEGMVAALAGLQGLRELELA
SNHFYLPRDLVAQPSLKHLDLNNLSVLTYSFRNTHLESLEHLEDAKLVKING
TLAELQGHPRHIVLDDNNPWCDCYVFLGIVLALIGALIFLLVLYLNKGIKKWMH
LELNSADLDCDLPSPQTSYVFLGIVLALIGALIFLLVLYLNKGIKKWMH
RDHMEGYHYRYEINADPRLTNLSSSDV"

ORIGIN

Alignment Scores:
Pred. No.: 16.7 Length: 2379
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x BC037161 (1-2379)

Qy 1 ValLeuTyLeuAsnArgLyGlylle 9
Db 1543 GTTTTGTATTGACCGCAAGGGGATA 1569

RESULT 31
BC058198
LOCUS
DEFINITION BC058198 Mus musculus trophoblast Glycoprotein, mRNA (CDNA clone MGC:68145
IMAGE:5353871), complete cds.
ACCESSION BC058198
VERSION BC058198.1 GI:34849573
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 2423)
Strausberg,R.B., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,C.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shvachenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalish,D.E.,
Schnerker,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2423)
Strausberg,R.
Direct Submission
Submitted (15-SEP-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,K., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Taurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 123 Row: p Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755854.
FEATURES
Location/Qualifiers
1..2423
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:68145 IMAGE:5353871"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..2423
/gene="tpbg"
/note="synonym: 574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
402..1682
/gene="tpbg"
/codon_start=1
/product="Tpbg protein"
/protein_id="AAH58198.1"
/db_xref="GI:34849574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
/translation="MPGAGRGPAGDRLRLARLALVLLGVWSSSPSSPTSSSPSS
PFLVAGAPARQPLAEALAAALNSGRLSDVDEVRAGAFELPSRLQDLNPLADLSPF
AFAGSNASVSPLEELLNHIVPPEDORQNGSPGMAFEGMVAALAGLQGLRELELA
TCLEASNHFLPRLDLAQLPSLKHLDLNNLSVLTYSFRNTHLESLEHLEDAKLVKING
TLAELQGHPRHIVLDDNNPWCDCYVFLGIVLALIGALIFLLVLYLNKGIKKWMH
MNRGLLNSDLDCAVLPQSLQTSYVFLGIVLALIGALIFLLVLYLNKGIKKWMH
NRDRCRDMEGYHYRYEINADPRLTNLSSSDV"
642..1262
/gene="Tpbg"
/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [function unknown]"
/db_xref="CDD:COG4886"
1299..1415
/gene="Tpbg"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

```

Alignment Scores:
Pred. No.: 17 Length: 2423
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-14 (1-9) x BC058198 (1-2423)

Qy 1 ValLeuTyTLauAsnArgLysGlyIle 9
Db 1536 GTTTTGTATTGACCGTAAGGCATA 1562

RESULT 32
LOCUS AX961912 2557 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 123 from Patent WO03104277.
ACCESSION AX961912
VERSION AX961912.1 GI:40881322
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Sugahara, T., Matsuura, A., Honda, G., Muramatsu, S. and Ishizawa, K.
Stat6 activation gene
Patent: WO 03104277-A 123 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
Location/Qualifiers
1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881322"
/translation="MPGAGSRGSPAGDGLRLRLALVLLGWVSASAPSSVPSSSTS
PADFLASGAQPPPAERCPAAECSEAAATVKVNRNLLVPAADLPVYRNLFITGNQ
MTVLPAAGAPQPLADLEALNLSGNHLKEVCAGAFELPLGLRLDLSHNPITNLSAF
VPAGSNASVSPLEELILNHI VPPEDORONGSPGVMVAFEGMAAALRSLALRGL
TRLEASNHFLPLRDLIAQLPSLAYLDLRNNSLYSLTYASFRLNLTLSLHLEDNAL
KVLHNSHAEWGGLAHVKVFLDNNPWCDCYADNADVAMVAMKETEVPDKARLTCAPEK
MRNRGLDLSDDLDCDAVLPSQTSYVFLGIVLALIGAIFLLVLYLNKRGKIKMWH
NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN
Alignment Scores:
Pred. No.: 18 Length: 2557
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-14 (1-9) x AX961914 (1-2557)

Qy 1 ValLeuTyTLauAsnArgLysGlyIle 9
Db 1690 GTTTTGTATTGACCGTAAGGCATA 1716

RESULT 34
LOCUS AB168308 2714 bp mRNA linear PRI 18-JUN-2005
DEFINITION Macaca fascicularis testis cDNA clone: Q5a-11109, similar to human
trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.
ACCESSION AB168308
VERSION AB168308.1 GI:67967899
KEYWORDS oligo capping; fib (full insert sequence).
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE
1 International consortium for macaque cDNA sequencing and analysis.
DNA sequences of macaque genes expressed in brain or testis and its
evolutionary implications
Unpublished

JOURNAL
REFERENCE
AUTHORS Osada, N., Hirata, M., Tanuma, R., Kusuda, J., Hida, M., Suzuki, Y.,
Sugano, S., Gojobori, T., Shen, J.-C.-K., Wu, C.-I. and Hashimoto, K.
TITLE Substitution rate and structural divergence of 5'UTR evolution:
Comparative analysis between human and cynomolgus monkey cDNAs
Unpublished
JOURNAL
REFERENCE
AUTHORS Hashimoto, K., Kusuda, J. and Sugano, S.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-2004) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan

```

(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/, Tel: 81-3-5285-1111 (ex: 2120), Fax: 81-3-5285-1181)
 The international consortium for macaque cDNA sequencing and analysis consists of: Department of Virology and Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; Division of Genetic Resources, National Institute of Infectious Diseases of Japan, Tokyo, Japan; National Health Research Institute, Taipei, Taiwan; Institute of Molecular Biology, Academia Sinica, Taipei, Taiwan; Department of Ecology & Evolution, University of Chicago, Chicago, IL, USA; Center for Information Biology, National Institute of Genetics of Japan, Mishima, Japan.
 Clone distribution: clone distribution information can be found at: http://www.nih.go.jp/yoken/genebank/

Lab host: TOP10

Vector: pME18S-FL3 (Acc.NO. AB009864)

R. Site1: DraIII (CAGCTGTG)

R. Site2: DraIII (CAGCATGTG)

Description: 1st strand cDNA was primed with an oligo(dT) primer using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method. Libraries were made from:

Qcc8: cerebellum cortex

QnpA: parietal lobe

QtrA: temporal lobe right

QflA: frontal lobe left

QmoA: medulla oblongata

QbsA: brain stem

QorA: occipital lobe right

QtsA: testis

Custom primers were used for 5' and 3'-end sequencing. The full-insert sequencing was done by primer-walking method using ABI DNA sequencer.

FEATURES

source

Location/Qualifiers

1..2714

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="QtsA-11109"

/sex="male"

/clone_lib="macaque cDNA library QtsA"

/dev_stage="adult"

764..2026

/note="unnamed protein product; Homo sapiens trophoblast

glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3"

/codon_start=1

/protein_id="BAE00432.1"

/db_xref="GI:67967900"

/translation="MPGCGSRGPAAGDGRRLRLARLALVLVWSSSSSTSSASSSSSS

APFLASAVSQPPLDQCPCAESEAAATKVCNRLNTEVPTDLPYVYRNFLTGNQ

LAVLPAGAFARPPPLAEALNLGSELDEVRAGAFELPSLRQLDLSHNPLADLSPP

AFSGNASVSAPSLVELLNHI VPPDDRQNRSGFEGVAAALVAGRALQGLRLLELA

SNHFLYLRDVLQALPSRLYLDLNNSLVSLTVSPRNLTHLESILHLENALKVLHNG

TLAELQGLPHRVFLDNPWCDCHMADVTWLKQTVVGQDRLTCAYPEKRNRYL

LELNSADLDCDLPSPSLQTSYVFLGIVLALIGAILFLVLVLRNGIKKKWMHNRDAC

RDHMEGYHYRYEINADPRLTNLSSSDV"

ORIGIN

Alignment Scores:			
Pred. No.:	19.1	Length:	2714
Score:	45.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-774-176-14 (1-9) x AB168308 (1-2714)

QY 1 ValLeuTyrLeuAsnArgLysGlyIle 9

|||||

Db 1880 GTTTGCTATTGACCGCAAGGGGATA 1506

RESULT 35

HSA012159

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

misc_binding

misc_binding

gene

exon

intron

exon

CDS

sig_peptide

mat_peptide

polyA_signal

1..5551

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

2698..2703

/bound_moiety="Sp1"

2704..2709

/bound_moiety="Sp1"

2716..5400

/gene="574"

/evidence="experimental"

2801..3092

/gene="574"

/evidence="experimental"

3093..5400

/gene="574"

/evidence="experimental"

3431..4693

/gene="574"

/codon_start=1

/product="574 oncofetal trophoblast glycoprotein"

/protein_id="CAA09930.1"

/db_xref="GI:3805947"

/db_xref="GOA:Q13641"

/db_xref="InterPro:IPR000372"

/db_xref="InterPro:IPR000483"

/db_xref="InterPro:IPR001611"

/db_xref="InterPro:IPR003591"

/db_xref="UniProt/TrEMBL:Q13641"

/translation="MPGCGSRGPAAGDGRRLRLARLALVLVWSSSSSTSSASSSSSS

APFLASAVSQPPLDQCPCAESEAAATKVCNRLNTEVPTDLPYVYRNFLTGNQ

LAVLPAGAFARPPPLAEALNLGSELDEVRAGAFELPSLRQLDLSHNPLADLSPP

AFSGNASVSAPSLVELLNHI VPPDDRQNRSGFEGVAAALVAGRALQGLRLLELA

SNHFLYLRDVLQALPSRLYLDLNNSLVSLTVSPRNLTHLESILHLENALKVLHNG

TLAELQGLPHRVFLDNPWCDCHMADVTWLKQTVVGQDRLTCAYPEKRNRYL

LELNSADLDCDLPSPSLQTSYVFLGIVLALIGAILFLVLVLRNGIKKKWMHNRDAC

RDHMEGYHYRYEINADPRLTNLSSSDV"

3431..3516

/gene="574"

3517..4690

/gene="574"

/product="574 oncofetal trophoblast glycoprotein"

5331..5336

```

/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProt/TREMBL:Q920L0"
/translation="MPGAGSGPSAGDGRRLRLARLALLVLLGWVSASAPSSVPSSTS
PADFLAGSAQPPPAERCPAAECSEAAITVCVARNLLLEVADLPYPVYNLPLTGNQ
MTVLPAAGAPAROPPLADLEALNLGNHLKEVCAGAFELHPLGLRLDLNPLNLTNSAF
VPAGSNASVAPSPLLEELIANHIVPEDORONGSPGMYAFSGMYAALRSGIALRGL
TKLEASNHFLPLPRDLALQLPSLYLRLNNLSVLTYSFANLTHLSLHLEDNAL
KYLHNSTLAEWQGLAHVKVFLDNNPWCDCTMADVMWLKETEVEVDPKARLTCAFFEK
MENRGLDLSGLDCLDAVLQSLQTSYVFLGIVLALIGAILFLVLYLNKRGIKKMMH
NIRDACRDHMEGYHYRINEADPLRLTSLSSNDV"
3779..3865
/gene="574"
3866..5056
/gene="574"
/product="574 oncofetal trophoblast glycoprotein"
5713..5718
/gene="574"
5759..5764
/gene="574"

sig_peptide
mat_peptide

polyA_signal
polyA_signal

ORIGIN
Alignment Scores:
Pred. No.: 56.1 Length: 7942
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0
US-10-774-176-14 (1-9) x MMU012160 (1-7942)
QY 1 ValLeuTyLeuAsnArgLysGlyIle 9
|||||
Db 4913 GTTTGTTATTTCAGCGTAAGGCATA 4939

RESULT 37
HSJ492P14
LOCUS
DEFINITION
HSJ492P14 121909 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone Rp3-492p14 on chromosome 6q13-15
Contains a single stranded DNA binding protein pseudogene, the TPBG
gene for trophoblast glycoprotein (574-AG) and a CpG island,
complete sequence.
ACCESSION
AL121977
VERSION
AL121977.11 GI:11863678
KEYWORDS
HTG; CpG island; TPBG.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 121909)
Garner, P.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Dec 15, 2000 this sequence version replaced gi:11558491.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Swi, SWISSPROT; Tr:, TREMBL; Wp:, WormPeP; Information
on the WormPeP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
Rp3-492p14 is from the library RPCI-3 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute

```

Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

Location/Qualifiers

1..121909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="q13-15"
/clone="RP3-492P14"
/clone_lib="RPCI-3"
100
/note="Clone right end: RP1-93K22"
complement(10004.._10982)
/locus_tag="RP3-492P14.2-001"
/pseudoc
complement(10004.._10982)
/locus_tag="RP3-492P14.2-001"
/note="match: proteins: P81877 Q99LX9 Q9BWW6 Q9CYZ8 Q9D6L4 Q9P038 Q9Y477"
/pseudoc
/codon_start=1
86539
/note="Clone left end: RP1-90G1"
109639.._116836
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
join(109639.._109916,110631.._116836)
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
/product="trophoblast glycoprotein"
/note="match: ESTs: AA143121 AA152323 AA565852 AA643734 AL544610 AW471072 AW662538 BE260089 BF306457 BF306926 BF314984 BI196133 BI562387 BM069633 BM670613
match: cDNAs: AJ420536.1 Z29083.1"
110970.._112232
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
/standard_names="OTTHUMP0000016786"
/note="match: proteins: Q13641 Q9QYD9 Q9Z0L0"
/codon_start=1
/product="trophoblast glycoprotein"
/protein_id="CAI21546.1"
/db_xref="GI:56203539"
/db_xref="Genew:12004"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000172"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCSRGPAAGDGRRLRLARLALLVLLGWSSSPSSASSPSSS
APFLASVAPPLDQPCALCESEARTVKVNRNLTEVPTDLPAYVRNLPLTGNQ
LAVLPAGAFARPPLAEALNALNSGRDLRDEVRAGAFELPSLRQLDLSNPLADLSPP
AFSGNSVSPSPVELLNHIVPEDERQNRSEFGMVVAALLAGRAQGGRLLELA
SNHFLYLRDLVAQLPSLRHLSDNNLSVLSVTFVRNLTLSLESLHEDNALKVLHG
TLAEQGLPHTRVFLDNNPWCDCMADMTWLKETEYVQKDRLTCTAYPERMRNL
LELNSADLDCDPLPSPQTSYVFLGIVLALICAILLLVLYLNKRGIKKWMHNRDAC
RDHMEGYHYRVEINADPRLTNLSSNDV"
116817.._116822
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
116836
/gene="TPBG"

polyA_signal

polyA_site

misc_feature
121909
/locus_tag="RP3-492P14.1-001"
/note="Clone_right_end: RP3-492P14"

ORIGIN

Alignment Scores:
Pred. No.: 871 Length: 121909
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-14 (1-9) x HSJ492P14 (1-121909)
QY 1 ValLeuTyxIeuAsnArgLysGlyIle 9
|||||
DB 112086 GTTTGTAATTGAACCGCAAGGGGATA 112112
|||||

RESULT 38
AC158516/c
LOCUS
DEFINITION Mus musculus BAC clone RP24-511A23 from chromosome 9, complete
sequence. 167046 bp DNA linear ROD 21-JUN-2005
AC158516 AC117768
ACCESSION
VERSION AC158516.2 GI:63025421
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
Adams.S., Cotton.M. and Haglund.K.
1 (bases 1 to 167046)
TITLE The sequence of Mus musculus BAC clone RP24-511A23
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 167046)
AUTHORS Wilson.R.K.
TITLE Direct Submission
JOURNAL Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 167046)
REFERENCE Wilson.R.K.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 167046)
REFERENCE Wilson.R.K.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On May 4, 2005 this sequence version replaced gi:61656412.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0511A23
Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e. phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. The assembly was confirmed by restriction digest.

This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal.

If a sequence meets the criteria of the above statement, it needs no comments or tags. If the criteria are not met, such as ambiguous bases, then the region is duly annotated.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

FEATURES

Location/Qualifiers
1..167046
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/clone="RP24-511A23"
/clone_lib="RPC1-24"
16685..16712
/note="Sequence derived from PCR product of genomic DNA"

misc_feature
unsure
unsure
unsure
ORIGIN
Alignment Scores:
Pred. No.: 1.2e+03 Length: 167046
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-14 (1-9) x AC158516 (1-167046)
Qy 1 ValLeuTyrLeuAsnArgIysGlyIle 9
|||||
Db 109704 GTTTTGTATTTCGACCGTAAGGCATA 109678
|||||

RESULT 39
AC128294/c AC128294 210237 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.
DEFINITION AC128294
ACCESSION AC128294
VERSION AC128294.3 GI:25083347
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 210237)
AUTHORS Munz, D. Marie., Metzker, M. Lee., Abranson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chaves, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Sacoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorovich, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louiseged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, T., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoileme, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Unpublished
Direct Submission
2 (bases 1 to 210237)
Worley, K. C.

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 210237)
Rat Genome Sequencing Consortium.

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265004. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine


```

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G2GV
Center clone name: CH230-176H20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 201781 bases at least Q40
Consensus quality: 203921 bases at least Q30
Consensus quality: 205310 bases at least Q20
Estimated insert size: 205531; sum-of-coverage: 7x in Q20 bases; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 210237: contig of 210237 bp in length.
Location/Qualifiers
1..210237
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-176H20"
1..1142
/feature="wgs_end_extension"
clone_end:77
2177..144799
/feature="clone_boundary"
clone_end:77
site:
end sequence: BH360464"
complement(206062..206961)
/feature="clone_boundary"
clone_end:5p6
site:
end sequence: BH360465"
208907..210237
/feature="wgs_end_extension"
clone_end:5p6"

Alignment Scores:
Pred. No.: 1.51e+03 Length: 210237
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0

ORIGIN
Alignment Scores:
Pred. No.: 1.51e+03 Length: 210237
Score: 45.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-14 (1-9) x AC128294 (1-210237)
Oy 1 ValLeuTyRLeuAsnArgLysGlyIle 9
|||||
Db 110384 GTTTTGATTGACCGGAGGCATA 110358

RESULT 40
AC106962/c
LOCUS Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4
DEFINITION Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4
ACCESSION AC106962
VERSION AC106962.5 GI:25139469
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)

```

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 239076)

Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Davidland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,K., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M., Hollins,B., Howells,S., Huiyik,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulsegged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L., L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlaczky,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE

Direct Submission

REFERENCE
AUTHORS

2 (bases 1 to 239076)

Worley,K.C.
Direct Submission
Submitted (14-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS

3 (bases 1 to 239076)

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 20, 2002 this sequence version replaced gi:22857070.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled with Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GOPI

Center clone name: CH230-87110

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 228642 bases at least Q40

Consensus quality: 232269 bases at least Q30

Consensus quality: 234041 bases at least Q20

Estimated insert size: 231522; sum-of-contigs estimation

Quality coverage: % in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 234710: contig of 234710 bp in length

* 234711 234810: gap of unknown length

* 234811 235924: contig of 1114 bp in length

* 235925 236024: gap of unknown length

* 236025 237314: contig of 1290 bp in length

* 237315 237414: gap of unknown length

* 237415 239076: contig of 1662 bp in length.

* Location/Qualifiers

1..239076

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-87110"

234711..234810

/estimated_length=unknown

235925..236024

/estimated_length=unknown

237315..237414

/estimated_length=unknown

ORIGIN

Alignment Scores:	1.71e+03	Length:	239076
Pred. No.:	45.00	Matches:	9
Score:	100.0%	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	100.0%	Indels:	0
Query Match:	14	Gaps:	0

US-10-774-176-14 (1-9) x AC106962 (1-239076)

Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9

Db 15595 GTTTTGTATTTCACCGGAAAGGCATA 15569

RESULT 41

AC131704/c

LOCUS

DEFINITION Mus musculus BAC clone RP24-149D9 from chromosome 12, complete

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

sequence.

AC131704

GI:30141995

HTG.

Mus musculus (house mouse)

Mus musculus

Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 163959)

Shah, N., Kozlowski, A. and Meyer, R.

The sequence of Mus musculus BAC clone RP24-149D9

Unpublished (2001)

2 (bases 1 to 163959)

Wilson, R.

Sequencing of Mus musculus

Unpublished (2001)

3 (bases 1 to 163959)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (25-AUG-2002)

Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 163959)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (11-FEB-2003)

Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 163959)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (26-APR-2003)

Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 163959)

Wilson, R.

Direct Submission

Submitted (27-NOV-2003)

Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Apr 26, 2003 this sequence version replaced gi:28302028.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0149D09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC122337

FEATURES
source

Location/Qualifiers
1..163959
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="12"
/map="12"
/clone="RP24-149D9"
/clone_lib="RPCI-24"
1..457
/rpt_family="L1"
720..993
/rpt_family="L1"
996..1119
/rpt_family="L1"
1120..1171
/rpt_family="ERV1"
1178..1454
/rpt_family="L1"
2110..2287
/rpt_family="B2"
4368..4433
/rpt_family="MER1_type"
4592..4684
/rpt_family="MIR"
6833..7744
/rpt_family="L1"
7825..8158
/rpt_family="L1"
8502..10277
/rpt_family="L1"
10265..11327
/rpt_family="L1"
14117..14409
/rpt_family="L1"
15909..16024
/rpt_family="B4"
16174..16253
/rpt_family="B4"
16700..16800
/rpt_family="ERVL"
16801..17133
/rpt_family="MaLR"
17134..17245
/rpt_family="ERVL"
17613..17825
/rpt_family="L1"
17941..18136
/rpt_family="L1"
19450..19764
/rpt_family="L1"
19968..20282
/rpt_family="MaLR"
20309..20676
/rpt_family="MaLR"
20765..21050
/rpt_family="L1"
21129..21261
/rpt_family="MaLR"
21348..21732
/rpt_family="MaLR"
22348..22455
/rpt_family="AcHobo"
23899..23956
/rpt_family="L2"
24516..24683
/rpt_family="B4"
24732..24983
/rpt_family="B4"
25597..25779
/rpt_family="B2"

repeat_region 26093..26642
/rpt_family="L1"
repeat_region 26632..26729
/rpt_family="L1"
repeat_region 26719..27640
/rpt_family="L1"
repeat_region 27700..27863
/rpt_family="MER1_type"
repeat_region 28670..29251
/rpt_family="L1"
repeat_region 29305..29872
/rpt_family="L1"
repeat_region 30175..31068
/rpt_family="L1"
repeat_region 31069..31457
/rpt_family="ERVK"
repeat_region 31458..31539
/rpt_family="L1"
repeat_region 32471..32797
/rpt_family="L1"
repeat_region 32903..33050
/rpt_family="B4"
repeat_region 33273..33462
/rpt_family="L1"
repeat_region 33625..33807
/rpt_family="L2"
repeat_region 33841..34293
/rpt_family="MaLR"
repeat_region 34243..35154
/rpt_family="ERVK"
repeat_region 35188..35506
/rpt_family="MaLR"
repeat_region 39035..39418
/rpt_family="MaLR"
repeat_region 39792..39932
/rpt_family="Alu"
repeat_region 40071..40330
/rpt_family="B4"
repeat_region 41548..41609
/rpt_family="L1"
unsure 42826..43032
/note="Unresolved simple sequence repeat."
repeat_region 43036..43182
/rpt_family="Alu"
repeat_region 45019..46853
/rpt_family="L1"
repeat_region 47139..47456
/rpt_family="L1"
repeat_region 47461..47699
/rpt_family="RMER1B"
repeat_region 48064..48408
/rpt_family="MaLR"

Alignment Scores:
Pred. No.: 3.21e+03 Length: 163959
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 95.6% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-14 (1-9) x AC131704 (1-163959)

QY 1 ValLeuTyRLeuAsnArgLysGlyFile 9
:::|||||
Db 88841 ATTTATATTAAATAGAAAGGGGTA 88815

RESULT 42
CQ736619
LOCUS
DEFINITION
ACCESSION
VERSION

CQ736619 615 bp DNA
Sequence 22553 from Patent WO02068579.
CQ736619
CQ736619.1 GI:42332470
linear PAT 03-FEB-2004

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
DEFINITION Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ACCESSION Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
VERSION Hominiidae; Homo.
KEYWORDS
SOURCE
ORGANISM
DEFINITION Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
ACCESSION Kits, such as nucleic acid arrays, comprising a majority of
VERSION humanexons or transcripts, for detecting expression and other uses
KEYWORDS thereof
SOURCE Patent: WO 02068579-A 22553 06-SEP-2002;
ORGANISM PE Corporation (NY) (US)
DEFINITION Location/Qualifiers
ACCESSION 1. 615
VERSION /organism="Homo sapiens"
KEYWORDS /mol_type="unassigned DNA"
SOURCE /db_xref="taxon:9606"
ORGANISM
DEFINITION Alignment Scores:
ACCESSION Pred. No.: 19.4 Length: 615
VERSION Score: 42.00 Matches: 8
KEYWORDS Percent Similarity: 100.0% Conservative: 1
SOURCE Best Local Similarity: 88.9% Mismatches: 0
ORGANISM Query Match: 93.3% Indels: 0
DEFINITION DB: 0
SOURCE Gaps: 0
US-10-774-176-14 (1-9) x CQ5736619 (1-615)
QY 1 ValLeuTyRLeuAsnArgLysGlycTyle 9
Db 499 GTGCTCTACCTAAACCGCGCGGCATC 525
RESULT 43
LOCUS CQ587318 4289 bp DNA linear PAT 02-FEB-2004
DEFINITION Sequence 15076 from Patent WO0171042.
ACCESSION CQ587318
VERSION CQ587318.1 GI:41646858
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
DEFINITION Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
ACCESSION Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
VERSION Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
KEYWORDS Ephydroidea; Drosophilidae; Drosophila.
SOURCE
REFERENCE Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
AUTHORS Detection kits, such as nucleic acid arrays, for detecting the
TITLE expression of 10,000 or more Drosophila genes and uses thereof
JOURNAL Patent: WO 0171042-A 15076 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1. 4289
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"
ORIGIN
Alignment Scores:
Pred. No.: 137 Length: 4289
Score: 42.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 93.3% Indels: 0
DB: 0
Gaps: 0
US-10-774-176-14 (1-9) x CQ587318 (1-4289)
QY 1 ValLeuTyRLeuAsnArgLysGlycTyle 9
Db 3275 ATCATATATCTTAAACCGTAAGGCATC 3301

RESULT 44
LOCUS AX254497 4289 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 64 from Patent WO0170980.
ACCESSION AX254497
VERSION AX254497.1 GI:16074220
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
DEFINITION Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
ACCESSION Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
VERSION Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
KEYWORDS Ephydroidea; Drosophilidae; Drosophila.
SOURCE
REFERENCE Cravchik, A.
AUTHORS Isolated g-protein coupled receptors, nucleic acid molecules
TITLE encoding gpcr proteins, and uses thereof as insecticidal targets
JOURNAL Patent: WO 0170980-A 64 27-SEP-2001;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1. 4289
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"
ORIGIN
Alignment Scores:
Pred. No.: 137 Length: 4289
Score: 42.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 93.3% Indels: 0
DB: 0
Gaps: 0
US-10-774-176-14 (1-9) x AX254497 (1-4289)
QY 1 ValLeuTyRLeuAsnArgLysGlycTyle 9
Db 3275 ATCATATATCTTAAACCGTAAGGCATC 3301
RESULT 45
LOCUS AC018122 10066 bp DNA linear HTG 09-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION AC018122
VERSION AC018122.1 GI:6553069
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
DEFINITION Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
ACCESSION Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
VERSION Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
KEYWORDS Ephydroidea; Drosophilidae; Drosophila.
SOURCE
REFERENCE 1 (bases 1 to 10066)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDW:10213448 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source Location/Qualifiers
1. 10066
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
ORIGIN
Alignment Scores:
Pred. No.: 322 Length: 10066
Score: 42.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2

Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 93.3% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-14 (1-9) x AC018122 (1-10066)

QY 1 ValLeuTyrlleuAnArglyvsGlylle 9
 DB 9051 ATCATATATCTTAAACGTAAGGCATC 9077

RESULT 46

AC079003/c

LOCUS AC079003 64168 bp DNA linear HTG 06-JAN-2001
 DEFINITION Homo sapiens chromosome 4 clone RP11-124C9 map 4, LOW-PASS SEQUENCE
 SAMPLING.

ACCESSION AC079003

VERSION AC079003.2 GI:12043592

KEYWORDS HTG; HTGS_PHASE0.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE 1 (bases 1 to 64168)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abrahams,H., Allen,N.,

TITLE Homo sapiens chromosome 4, clone RP11-124C9

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 64168)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abrahams,H., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bieda,F., Boguslavsky,L.,

Bouckgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,

Choepei,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,

DeAcellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,

FitzHugh,W., Gage,D., Galagan,J., Gardina,S., Ginde,S., Goyette,M.,

Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,

Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,

Lamazares,R., Landers,T., Lechoczky,J., Levine,R., Lieu,C., Liu,G.,

Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,

McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,

Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,

O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,

Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,

Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,

Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Tirrell,A., Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo,A.,

Wilson,B., Wu,X., Wyman,D., Ye.W.J., Young,G., Zainoun,J.,

Zimmer,A. and Zody,M.

TITLE Direct Submision

JOURNAL

COMMENT

Submitted (15-AUG-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jan 6, 2001 this sequence version replaced gi:9802826.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L10494

Center clone name: 124_C_9

* NOTE: This record contains 82 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that

* the record is updated, the accession number will
 * be preserved.

* 1 677: contig of 677 bp in length

* 678 777: gap of 100 bp

* 778 1447: contig of 670 bp in length

* 1448 1547: gap of 100 bp

* 1548 2248: contig of 701 bp in length

* 2249 2348: gap of 100 bp

* 3062: contig of 714 bp in length

* 3063 3162: gap of 100 bp

* 3163 3852: contig of 690 bp in length

* 3853 3952: gap of 100 bp

* 3953 4635: contig of 683 bp in length

* 4636 4735: gap of 100 bp

* 4736 5418: contig of 683 bp in length

* 5419 5518: gap of 100 bp

* 5519 6195: contig of 677 bp in length

* 6196 6295: gap of 100 bp

* 6296 6974: contig of 679 bp in length

* 6975 7074: gap of 100 bp

* 7075 7775: contig of 701 bp in length

* 7776 7875: gap of 100 bp

* 7876 8573: contig of 698 bp in length

* 8574 9353: gap of 100 bp

* 9354 10169: contig of 716 bp in length

* 9454 10269: gap of 100 bp

* 10170 10946: contig of 677 bp in length

* 10947 11046: gap of 100 bp

* 11047 11720: contig of 674 bp in length

* 11721 11820: gap of 100 bp

* 11821 12505: contig of 685 bp in length

* 12506 12605: gap of 100 bp

* 12606 13291: contig of 686 bp in length

* 13292 13391: gap of 100 bp

* 13392 14070: contig of 679 bp in length

* 14071 14170: gap of 100 bp

* 14171 14850: contig of 680 bp in length

* 14851 14950: gap of 100 bp

* 14951 15617: contig of 667 bp in length

* 15618 15717: gap of 100 bp

* 15718 16387: contig of 670 bp in length

* 16388 16487: gap of 100 bp

* 16488 17155: contig of 678 bp in length

* 17156 17265: gap of 100 bp

* 17266 17968: contig of 703 bp in length

* 17969 18068: gap of 100 bp

* 18069 18754: contig of 686 bp in length

* 18755 18854: gap of 100 bp

* 18855 19504: contig of 650 bp in length

* 19505 19604: gap of 100 bp

* 19605 20307: contig of 703 bp in length

* 20308 20407: gap of 100 bp

* 20408 21092: contig of 685 bp in length

* 21093 21192: gap of 100 bp

* 21193 21860: contig of 668 bp in length

* 21861 21960: gap of 100 bp

* 21961 22638: contig of 678 bp in length

* 22639 22738: gap of 100 bp

* 22739 23426: contig of 688 bp in length

* 23427 23526: gap of 100 bp

* 23527 24221: contig of 695 bp in length

* 24222 24321: gap of 100 bp

* 24322 25008: contig of 687 bp in length

* 25009 25108: gap of 100 bp

* 25109 25757: contig of 649 bp in length

* 25758 25857: gap of 100 bp

* 25858 26555: contig of 698 bp in length

* 26556 26655: gap of 100 bp

* 26656 27343: contig of 688 bp in length

* 27344 27443: gap of 100 bp

* 27444 28136: contig of 693 bp in length

```

* 28137 28236: gap of 100 bp
* 28237 28940: contig of 704 bp in length
* 28941 29040: gap of 100 bp
* 29041 29728: contig of 688 bp in length
* 29729 29828: gap of 100 bp
* 29829 30521: contig of 693 bp in length
* 30522 30621: gap of 100 bp
* 30622 31291: contig of 670 bp in length
* 31292 31391: gap of 100 bp
* 31392 32071: contig of 680 bp in length
* 32072 32171: gap of 100 bp
* 32172 32853: contig of 682 bp in length
* 32854 32953: gap of 100 bp
* 32954 33634: contig of 681 bp in length
* 33635 33734: gap of 100 bp
* 33735 34411: contig of 677 bp in length
* 34412 34511: gap of 100 bp
* 34512 35195: contig of 684 bp in length
* 35196 35295: gap of 100 bp
* 35296 35947: contig of 652 bp in length
* 35948 36047: gap of 100 bp
* 36048 36744: contig of 697 bp in length
* 36745 36844: gap of 100 bp
* 36845 37528: contig of 684 bp in length
* 37529 37628: gap of 100 bp
* 37629 38308: contig of 680 bp in length
* 38309 38408: gap of 100 bp
* 38409 39094: contig of 686 bp in length
* 39095 39194: gap of 100 bp
* 39195 39887: contig of 693 bp in length
* 39888 39987: gap of 100 bp
* 40073 40673: contig of 686 bp in length
* 40674 40773: gap of 100 bp
* 40774 41468: contig of 695 bp in length
* 41469 41568: gap of 100 bp
* 41569 42274: contig of 706 bp in length
* 42275 42374: gap of 100 bp
* 42375 43043: contig of 669 bp in length
* 43044 43143: gap of 100 bp
* 43144 43826: contig of 683 bp in length
* 43827 43926: gap of 100 bp
* 43927 44606: contig of 680 bp in length
* 44607 44706: gap of 100 bp
* 44707 45400: contig of 694 bp in length
* 45401 45500: gap of 100 bp
* 45501 46198: contig of 698 bp in length
* 46199 46298: gap of 100 bp
* 46299 46999: contig of 701 bp in length
* 47000 47099: gap of 100 bp
* 47100 47766: contig of 667 bp in length
* 47767 47866: gap of 100 bp
* 47867 48550: contig of 684 bp in length
* 48551 48650: gap of 100 bp
* 48651 49352: contig of 702 bp in length
* 49353 49452: gap of 100 bp
* 49453 50096: contig of 644 bp in length
* 50097 50196: gap of 100 bp
* 50197 50878: contig of 682 bp in length
* 50879 50979: gap of 100 bp
* 50979 51647: contig of 669 bp in length
* 51648 51747: gap of 100 bp
* 51748 52404: contig of 657 bp in length
* 52405 52505: gap of 100 bp
* 52505 53193: contig of 689 bp in length
* 53194 53293: gap of 100 bp
* 53294 53972: contig of 679 bp in length

Alignment Scores:
Pred. No.: 2.07e+03 Length: 64168
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.3% Indels: 0

```

```

DB: 14 Gaps: 0

US-10-774-176-14 (1-9) x AC079003 (1-64168)

QY 1 ValLeuTyLeuAsnArgLysGlyIle 9
Db 13817 GTGCTTACCTTAACCGCGCGGCATC 13791

RESULT 47
LOCUS AF188024/c
DEFINITION Homo sapiens chromosome 12 clone CTC-790L10 map 12p11.2, complete
sequence.
ACCESSION AF188024
VERSION AF188024.3 GI:23462852
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 111108)
AUTHORS Taudien,S., Wen,G., Schilhabel,M., Menzel,U., Jahn,N., Baumgart,C.,
Dette,M. and Rosenthal,A.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 111108)
AUTHORS Taudien,S., Wen,G.P., Schilhabel,M., Menzel,U., Jahn,N.,
Baumgart,C., Dette,M. and Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1999) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 111108)
AUTHORS Lagemann,D. and Platzer,M.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT On Oct 2, 2002 this sequence version replaced gi:14327767.

-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
-----
Project Information
Center project name: H269
Center clone name: CTC-790L10
-----
Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 110433 bases at least Q40
Consensus quality: 110887 bases at least Q30
Consensus quality: 111058 bases at least Q20
Quality coverage: 9.56x.

-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one pUC18 subclone; and the
assembly was confirmed by restriction digest

-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
Location/Qualifiers

```

FEATURES

```
source
1. .111108
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="12"
/maps="12p11.2, complete sequence."
/clone="CTC-790L10"
4532. .4740
/notes="single stranded/single chemistry region"
4601. .4604
/notes="low quality region , CTC-790L10"
4624
/notes="low quality region , CTC-790L10"
4695
/notes="low quality region , CTC-790L10"
4708
/notes="low quality region , CTC-790L10"
5188. .6336
/notes="single stranded/single chemistry region"
8356. .8544
/notes="single stranded/single chemistry region"
10370
/notes="low quality region , CTC-790L10"
10412
/notes="low quality region , CTC-790L10"
10596. .10846
/notes="single stranded/single chemistry region"
10625
/notes="low quality region , CTC-790L10"
10948
/notes="low quality region , CTC-790L10"
11045
/notes="low quality region , CTC-790L10"
11067. .11162
/notes="single clone coverage"
11073. .11110
/notes="low quality region , CTC-790L10"
11107. .11108
/notes="CTC-790L10"
11119. .11166
/notes="low quality region , CTC-790L10"
11144
/notes="CTC-790L10"
11157
/notes="CTC-790L10"
11159
/notes="CTC-790L10"
11173. .11224
/notes="low quality region , CTC-790L10"
11230
/notes="low quality region , CTC-790L10"
16055. .16095
/notes="single stranded/single chemistry region"
25625. .25634
/notes="low quality region , CTC-790L10"
47453. .47633
/notes="single stranded/single chemistry region"
50385. .50528
/notes="single stranded/single chemistry region"
53595
/notes="low quality region , CTC-790L10"
53840. .54199
/notes="single stranded/single chemistry region"
54011
/notes="low quality region , CTC-790L10"
54661. .55033
/notes="single stranded/single chemistry region"
54672
/notes="low quality region , CTC-790L10"
55034. .55045
/notes="single clone coverage"
55214. .55401
/notes="single stranded/single chemistry region"
```

```
misc_feature 61971. .62750
/notes="single stranded/single chemistry region"
misc_feature 62145. .62147
/notes="low quality region , CTC-790L10"
misc_feature 62253. .62257
/notes="low quality region , CTC-790L10"
misc_feature 63540. .64369
/notes="single stranded/single chemistry region"
misc_feature 70831. .71004
/notes="single stranded/single chemistry region"
misc_feature 72109. .72769
/notes="single stranded/single chemistry region"
misc_feature 72127
/notes="low quality region , CTC-790L10"
misc_feature 72160. .72161
/notes="low quality region , CTC-790L10"
misc_feature 73635. .73637
/notes="low quality region , CTC-790L10"
misc_feature 73669. .73672
/notes="low quality region , CTC-790L10"
misc_feature 73686. .73742
/notes="low quality region , CTC-790L10"
misc_feature 82541. .82560
/notes="low quality region , CTC-790L10"
misc_feature 85089
/notes="low quality region , CTC-790L10"
misc_feature 85156. .85899
/notes="single stranded/single chemistry region"
misc_feature 85181. .85187
/notes="low quality region , CTC-790L10"
misc_feature 85815. .85821
/notes="low quality region , CTC-790L10"
misc_feature 85856. .85871
/notes="low quality region , CTC-790L10"
misc_feature 85910
/notes="low quality region , CTC-790L10"
misc_feature 85959
/notes="low quality region , CTC-790L10"
misc_feature 110023. .110510
/notes="pcr product sequence only , CTC-790L10"
misc_feature 110511. .111108
/notes="pcr product sequence only , CTC-790L10"
misc_feature 110516. .110518
/notes="low quality region , CTC-790L10"
misc_feature 110527. .110546
/notes="low quality region , CTC-790L10"
misc_feature 110560. .110582
/notes="low quality region , CTC-790L10"
```

ORIGIN

```
Alignment Scores: 3.59e+03 Length: 111108
Pred. No.: 42.00 Matches: 8
Score: 100.0% Conservative: 1
Percent Similarity: 88.9% Mismatches: 0
Best Local Similarity: 93.3% Indels: 0
Query Match: 8 Gaps: 0
DB:
```

US-10-774-176-14 (1-9) x AF188024 (1-111108)

```
QY 1 ValLeuTyRLeuAsnArgLysGlyIle 9
|||||:|||||:|||||:|||||:|||||:
Db 59580 GTRACTGTTTAAATAGAAAGGANYA 59554
```

RESULT 48

AC022363/c

LOCUS

DEFINITION

AC022363

ACCESSION

VERSION

KEYWORDS

SOURCE

114933 bp DNA linear PRI 21-AUG-2001
Homo sapiens 12 BAC RP11-1151B7 (Roswell Park Cancer Institute
Human BAC library) complete sequence.
AC022363
AC022363.24 GI:11067115
Homo sapiens (human)


```

#      350
bases 300
      250
      200
      150
      100
       50
       0
-----
      5  10  15  20  25  30  35  40
Phrap Value Range

```

Version: 1.01 qxf0.

FEATURES

source

Location/Qualifiers

```

1. .114933
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-1151B7"

```

misc_feature

1. .205

```

/note="Overlaps bases 73353. .73557 of clone AC022364"
/function="Overlaps with adjacent clone AC022364"
complement(2174. .3412)
/rpt_family="L1PAL6"

```

repeat_region

3566. .3605

/rpt_family="AT_rich"

repeat_region

4696. .4718

/rpt_family="AT_rich"

repeat_region

5165. .5349

/rpt_family="MER91B"

repeat_region

5651. .5700

/rpt_family="L2"

repeat_region

7016. .7067

/rpt_family="AT_rich"

repeat_region

7362. .7782

/rpt_family="MSTA"

repeat_region

7951. .8090

/rpt_family="MIR"

repeat_region

complement(8162. .8985)

/rpt_family="L1PA3"

repeat_region

complement(9047. .9198)

/rpt_family="MER5B"

repeat_region

11253. .11618

/rpt_family="THE1B"

Alignment Scores:

```

Pred. No.: 3.71e+03 Length: 114933
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.3% Indels: 0
DB: 8 Gaps: 0

```

US-10-774-176-14 (1-9) x AC022363 (1-114933)

Qy 1 ValLeuTyfLeuAsnArgLysGlyIle 9

Db 67686 GTACTGTATTAAATAAGAGAGGAATA 67660

RESULT 49

AC012164/c

LOCUS

DEFINITION Drosophila melanogaster clone BACR14D22, complete sequence.

AC012164

VERSION

AC012164.12

KEYWORDS

HTG.

SOURCE

Drosophila melanogaster (fruit fly)

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE

AUTHORS

Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 169385)
 Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
 Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
 Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
 Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
 Kim,B., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
 Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
 Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
 Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
 Rubin,G.M.

TITLE

JOURNAL

Direct Submission
 Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA

REFERENCE

AUTHORS

2 (bases 1 to 169385)
 Celniker,S., Carlson,J., Wan,K., Pfeiffer,B., Frise,E., George,R.,
 Hoskins,R., Stapleton,M., Pacleb,J., Park,S., Svirskas,R.,
 Smith,E., Yu,C. and Rubin,G.

TITLE

JOURNAL

Direct Submission
 Submitted (31-JUL-2004) Berkeley Drosophila Genome Project, MS
 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, US

COMMENT

On Jul 31, 2004 this sequence version replaced gi:13324738.
 Sequence submitted by:
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This sequence submission incorporates changes made during
 reevaluation of the assembly or fingerprint verification of the
 clone. For further information about this sequence, including its
 location and relationship to other sequences, please visit our
 sequence archive web site (<http://www.fruitfly.org/sequence/>) or
 send email to bdgp@fruitfly.org.

FEATURES

source

1. .169385

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/strain="y; cn bw sp"

/db_xref="taxon:7227"

/chromosome="X"

/map="17C-17B"

/clone="BAC clone BACR14D22 (D1120)"

/clone_lib="RPC1-98 (Roswell Park Cancer Institute
 Drosophila melanogaster BAC library, partial EcoRI in
 pBACs3.6)"

ORIGIN

```

Alignment Scores:
Pred. No.: 5.48e+03 Length: 169385
Score: 42.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 93.3% Indels: 0
DB: 2 Gaps: 0

```

US-10-774-176-14 (1-9) x AC012164 (1-169385)

Qy 1 ValLeuTyfLeuAsnArgLysGlyIle 9

Db 63180 ATCATATATCTTAACCGTAAGGCATC 63154

RESULT 50

AC022045

LOCUS

DEFINITION Homo sapiens clone RP11-11G5, WORKING DRAFT SEQUENCE, 6 unordered
 pieces.

AC022045

VERSION

AC022045.2

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.
1 (bases 1 to 175837)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-11G5
Unpublished
2 (bases 1 to 175837)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
Choepe,V., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fensetor,J.,
Farreira,P., FitzHugh,M., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K.,
McDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6751775.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3163
Center clone name: 11 G 5
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 172439 bases at least Q40
Consensus quality: 174228 bases at least Q30
Consensus quality: 174815 bases at least Q20
Insert size: 174000; agarose-fp
Insert size: 175337; sum-of-contigs
Quality coverage: 5.9 in Q20 bases; agarose-fp
Quality coverage: 5.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 5925: contig of 5925 bp in length
* 5926 6025: gap of 100 bp
* 6026 21447: contig of 15422 bp in length
* 21448 21547: gap of 100 bp
* 21548 50643: contig of 29096 bp in length
* 50644 50743: gap of 100 bp
* 50744 81519: contig of 30776 bp in length
* 81520 127107: gap of 100 bp
* 127108 127207: contig of 45488 bp in length
* 127208 175837: contig of 48630 bp in length.
Location/Qualifiers
1..175837
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="taxon:9606"
/clone_lib="RPC1-11 Human Male BAC"
1..5925
/note="assembly_fragment"
5926 6025
/estimated_length=100
6026..21447
/note="assembly_fragment"
21448..21547
/estimated_length=100
21548..50643
/note="assembly_fragment"
50644..50743
/estimated_length=100
50744..81519
/note="assembly_fragment"
clone_end:77
vector_side:left
81520..81619
/estimated_length=100
81620..127107
/note="assembly_fragment"
127108..127207
/estimated_length=100
127208..175837
/note="assembly_fragment"
clone_end:SP6
vector_side:left

ORIGIN
Alignment Scores:
Pred. No.: 5.69e+03 Length: 175837
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.3% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-14 (1-9) x AC022045 (1-175837)
Qy 1 ValLeuTyLeuAsnArgLysGlyIle 9
Db 94401 GTGCTTACCTAAACCGCGGCATC 94427

Search completed: April 25, 2006, 20:38:19
JOB time : 3079.7 secs

FEATURES
source

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds
(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-13

Perfect score: 47

Sequence: 1 FLYLPRDL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/absas/ABSSWEB.spool/US1074176/runat_24042006_165112_19185/app_query.fasta.1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cd1 -LIST=1000
-DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05p
-USR=US1074176 @CGN 1.1 3463 @runat_24042006_165112_19185 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 21.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001as.*

5: Geneseq2001bs.*

6: Geneseq2002as.*

7: Geneseq2002bs.*

8: Geneseq2003as.*

9: Geneseq2003bs.*

10: Geneseq2003cs.*

11: Geneseq2003ds.*

12: Geneseq2004as.*

13: Geneseq2004bs.*

14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	927	6	ABT07721 Breast ca
2	47	100.0	927	8	ABX76333 Lung canc
3	47	100.0	927	10	ADB80503 Ovarian c
4	47	100.0	927	11	ADN38723 Cancer/an

5	47	100.0	973	8	AAD56198
6	47	100.0	1156	6	ABV99349
7	47	100.0	1263	3	AAA27058
8	47	100.0	1263	4	AAF89736
9	47	100.0	1263	6	ABK87174
10	47	100.0	1331	8	AAD56199
11	47	100.0	2020	10	ADJ56299
12	47	100.0	2053	8	ACC51052
13	47	100.0	2053	8	ABX76332
14	47	100.0	2053	8	AAD56197
15	47	100.0	2053	8	AAD56200
16	47	100.0	2053	11	ADN38721
17	47	100.0	2053	12	ADL06473
18	47	100.0	2053	12	ADN03961
19	47	100.0	2053	13	ADR25444
20	47	100.0	2053	13	ACN38510
21	47	100.0	2053	13	ADV35098
22	47	100.0	2338	5	AAS87175
23	47	100.0	2359	4	AAK94253
24	47	100.0	2359	12	ADL30831
25	47	100.0	2361	4	AAK94254
26	47	100.0	2361	12	ADL30833
27	47	100.0	2361	12	ADL30833
28	43	91.5	1260	6	ABK87175
29	43	91.5	1260	10	ADB97513
30	43	91.5	1260	10	ADB97452
31	43	91.5	86248	10	ADC00087
32	43	91.5	87563	9	ACD19044
33	41	87.2	901	3	AAA27060
34	40	85.1	1268	13	ADT16562
35	40	85.1	1281	3	AAA27059
36	40	85.1	2557	12	ADL26160
37	40	85.1	2557	12	ADL26158
38	40	85.1	2568	13	ADT45579
39	40	85.1	2571	13	ADS46521
40	40	85.1	34488	5	AAF97854
41	40	85.1	121162	3	AAC66548
42	39	83.0	108	2	AQ34023
43	39	83.0	584	13	ADQ49285
44	39	83.0	753	14	ABE53920
45	39	83.0	937	8	ACC62259
46	39	83.0	945	9	AAU57532
47	39	83.0	945	12	ADQ87521
48	39	83.0	1058	8	ACC62260
49	39	83.0	1165	14	ABE53915
50	39	83.0	1241	10	ADL40459
51	39	83.0	1280	14	ABE53911
52	39	83.0	1340	3	ADN00690
53	39	83.0	1346	8	ACC62258
54	39	83.0	1346	8	ACC62258
55	39	83.0	1358	4	AAK41256
56	39	83.0	1373	14	ADX06033
57	39	83.0	1374	14	ADV60518
58	39	83.0	1386	5	ABV23063
59	39	83.0	1386	5	ABV28899
60	39	83.0	1404	8	ACC50288
61	39	83.0	1431	6	ABT07729
62	39	83.0	1439	14	ABE53909
63	39	83.0	1712	11	ADM01708
64	39	83.0	1820	12	ADH13724
65	39	83.0	1957	14	ABE53913
66	39	83.0	1989	4	AAK94094
67	39	83.0	1989	12	ADL30527
68	39	83.0	2000	6	ABZ15552
69	39	83.0	2374	11	ADM03693
70	39	83.0	2561	14	ABE53921
71	39	83.0	2562	6	ABE54887
72	39	83.0	2678	5	ABV29320
73	39	83.0	2678	5	ABV23462
74	39	83.0	2679	11	ACN92098
75	39	83.0	3310	8	ABZ74454
76	39	83.0	3310	8	ADA98878
77	39	83.0	3310	10	ADC20907

78	39	83.0	3310	10	ABT16993	Abt16993 Human sec	C 151	37	78.7	99656	14	AEb32414	Human gen
79	39	83.0	3310	10	ABe68001	Abt68001 Human sec	C 152	37	78.7	99656	14	AEb32364	Human gen
80	39	83.0	3953	4	ABl07580	Drosophil	C 153	37	78.7	161280	10	ADb67054	Human lun
81	39	83.0	28676	4	AAk80349	Human imm	C 154	37	78.7	161280	10	ADb68308	Human lun
82	39	83.0	110000	12	ADQ34435_4	Continuation (5 of	C 155	36	76.6	25	6	ABV81322	Human HTP
83	38	80.9	516	4	AAK83470	Human imm	C 156	36	76.6	25	6	ABV81324	Human HTP
84	38	80.9	533	10	ADd48785	Rat gene	C 157	36	76.6	25	6	ABV81325	Human HTP
85	38	80.9	565	13	ADU09778	Solid tum	C 158	36	76.6	25	6	ABV81321	Human HTP
86	38	80.9	655	13	ADQ56661	Novel can	C 159	36	76.6	25	6	ABV81323	Human HTP
87	38	80.9	699	7	ADG31382	Human gen	C 160	36	76.6	60	6	ABN49974	Human sp1
88	38	80.9	699	7	ADY36770	HIRA geno	C 161	36	76.6	181	10	ADI19158	ADI19158 A. thalia
89	38	80.9	916	14	ABE45794	Tomato ne	C 162	36	76.6	290	5	AAH88106	Peppermin
90	38	80.9	2636	4	AAI61106	Human pol	C 163	36	76.6	294	2	AAZ13860	Human gen
91	38	80.9	2957	13	ADX51094	Plant ful	C 164	36	76.6	370	10	ADC72730	Sequence
92	38	80.9	7390	4	AAK71702	Human imm	C 165	36	76.6	380	9	ACH49194	Human leu
93	38	80.9	7495	13	ADR30815	Zebrafish	C 166	36	76.6	428	9	ACH18679	Human adu
94	38	80.9	7508	13	ADR30814	Zebrafish	C 167	36	76.6	438	9	ACH27683	Human adu
95	38	80.9	110000	2	AAK31990_06	Continuation (7 of	C 168	36	76.6	446	10	ABT41464	Toxicity
96	38	80.9	226475	9	AAU58279	Human tum	C 169	36	76.6	451	14	ADZ61457	Rat 93667
97	38	80.9	273254	3	AAAC81914	Chlamydia	C 170	36	76.6	500	6	ABV78773	Human tes
98	37	78.7	256	10	ADP80620	Leukaemia	C 171	36	76.6	528	14	ACL58330	Human col
99	37	78.7	458	13	AEA01321	Oryza spe	C 172	36	76.6	582	3	AAA82329	N. mening
100	37	78.7	585	5	ABV52144	Human pro	C 173	36	76.6	715	2	AAZ15443	Human gen
101	37	78.7	601	14	ABE34258	Human DNA	C 174	36	76.6	734	6	ABK33933	DNA encod
102	37	78.7	601	14	ABE34257	Human DNA	C 175	36	76.6	734	8	ACAI11722	Human lun
103	37	78.7	601	14	ABE34259	Human DNA	C 176	36	76.6	734	8	ACA02908	Lung canc
104	37	78.7	759	4	AAH08403	Human cDN	C 177	36	76.6	734	10	ADH46950	Human lun
105	37	78.7	818	12	ADN11287	Human pp2	C 178	36	76.6	734	13	ADJ20869	Human lun
106	37	78.7	981	8	ACC42695	Epoxide h	C 179	36	76.6	763	14	ADY65655	S. mansoni
107	37	78.7	981	13	ADU00287	Epoxide h	C 180	36	76.6	954	6	ABK87584	DNA encod
108	37	78.7	1173	3	AAAC3700	Arabidops	C 181	36	76.6	957	12	ADJ87104	Nucleotid
109	37	78.7	1173	3	ADN94765	A. thalia	C 182	36	76.6	957	12	ADM33336	Human bit
110	37	78.7	1220	13	ADX30923	Plant ful	C 183	36	76.6	957	14	ADW74433	Human bit
111	37	78.7	1281	3	AAAC51096	Arabidops	C 184	36	76.6	957	14	ADW74649	Human bit
112	37	78.7	1418	12	ADQ21129	Human car	C 185	36	76.6	957	14	ADW74651	Human bit
113	37	78.7	1508	13	ADQ38908	Human car	C 186	36	76.6	1166	6	ABV78769	Human tes
114	37	78.7	1544	13	ADQ38910	Human SNP	C 187	36	76.6	1287	8	ACA21840	Prokaryot
115	37	78.7	1551	6	ABN95857	Gene #235	C 188	36	76.6	1441	3	AAAC39040	Arabidops
116	37	78.7	1579	3	AAAC37020	Arabidops	C 189	36	76.6	1489	13	ADT17460	Plant CDN
117	37	78.7	1583	3	AAAC33682	Arabidops	C 190	36	76.6	1767	10	ADB63552	Human cDN
118	37	78.7	1583	3	AAAC54749	Arabidops	C 191	36	76.6	1944	6	ABV83552	Human tes
119	37	78.7	1751	9	ACD19210	E. coli 0	C 192	36	76.6	1944	6	ABV78765	Human tes
120	37	78.7	1761	13	ADCO1567	Enterohae	C 193	36	76.6	2021	6	ABV83550	Human tes
121	37	78.7	1769	13	ADT15609	Plant cDN	C 194	36	76.6	2021	6	ABV78763	Human tes
122	37	78.7	1916	14	ABE67712	Rice geno	C 195	36	76.6	2299	13	ADX28529	Plant ful
123	37	78.7	2278	3	AAZ89962	Corn ADA2	C 196	36	76.6	2304	6	ABV78762	Human tes
124	37	78.7	2521	4	ABL26024	Drosophil	C 197	36	76.6	2355	6	ABK86121	DNA encod
125	37	78.7	3118	10	ADF82023	Leukaemia	C 198	36	76.6	2363	13	ADT17080	Plant CDN
126	37	78.7	3118	13	ADP55470	Human PRO	C 199	36	76.6	2364	8	ACA27674	Prokaryot
127	37	78.7	3662	4	ABL26000	Drosophil	C 200	36	76.6	2385	6	ABZ11329	Human pol
128	37	78.7	3935	11	ACL29592	Rice abio	C 201	36	76.6	2385	12	ADM43847	Novel hum
129	37	78.7	4185	14	ADZ61878	Murine Ez	C 202	36	76.6	2463	8	ACA30798	Prokaryot
130	37	78.7	5519	4	AAAC36740	Human car	C 203	36	76.6	2463	14	ADW64975	C. pneumo
131	37	78.7	5519	4	AAAC35819	Human car	C 204	36	76.6	2529	12	ADQ64549	Novel hum
132	37	78.7	5519	10	ADAE47434	Human car	C 205	36	76.6	2744	13	ADQ64549	Novel hum
133	37	78.7	5519	10	ADAE46513	Human car	C 206	36	76.6	2744	13	ADX12788	Plant ful
134	37	78.7	5519	13	ADJ08852	Human car	C 207	36	76.6	2755	13	ADX28444	Plant ful
135	37	78.7	5519	13	ADJ07931	Human car	C 208	36	76.6	2756	12	ADX27934	Plant ful
136	37	78.7	5667	4	AAH15191	Human cDN	C 209	36	76.6	2760	2	AAE1883	Marek's d
137	37	78.7	7072	4	AAAL03300	Human rep	C 210	36	76.6	2786	13	ADX28813	Plant ful
138	37	78.7	7073	4	AAAL03300	Human rep	C 211	36	76.6	2786	13	ADX28557	Plant ful
139	37	78.7	11363	4	AAK80081	Human imm	C 212	36	76.6	2795	13	ADX15130	Plant ful
140	37	78.7	11366	4	AAK80079	Human imm	C 213	36	76.6	2806	13	ADX15159	Plant ful
141	37	78.7	11366	4	AAK80080	Human imm	C 214	36	76.6	2826	13	ADX28607	Plant ful
142	37	78.7	13584	14	ABE64840	Escherich	C 215	36	76.6	2834	13	ADX28791	Plant ful
143	37	78.7	42379	12	ADQ97660	Mouse can	C 216	36	76.6	2865	6	ABV78760	Human tes
144	37	78.7	92562	10	ADAC85284	Human ITK	C 217	36	76.6	3032	10	ADI21860	Novel hum
145	37	78.7	92563	9	ADA02804	Human ITK	C 218	36	76.6	3047	12	ADP72740	Renal tox
146	37	78.7	92563	10	ADB72542	Human ITK	C 219	36	76.6	3296	6	ABV78759	Rat cardi
147	37	78.7	92563	12	ADM74399	Human car	C 220	36	76.6	3296	6	ABV78761	Human tes
148	37	78.7	96389	9	ADW02675	Mouse Top	C 221	36	76.6	4048	4	ABL08264	Drosophil
149	37	78.7	96389	10	ADB72413	Mouse Top	C 222	36	76.6	4074	13	ADX14821	Plant ful
150	37	78.7	96389	10	ADB95923	Mouse Top	C 223	36	76.6	5009	5	AAH27470	DRC1 geno

224	36	76.6	5011	4	ABK43166	Abk43166 DNA encod	C 297	35	74.5	473	4	AAI21709	Aai21709 Probe #11
C 225	36	76.6	5162	10	ADB79936	Adb79936 Human put	C 298	35	74.5	473	4	AAI12518	Aai12518 Probe #24
C 226	36	76.6	5399	2	AAV37932	Aav37932 DNA encod	C 299	35	74.5	473	4	ABA54224	Abas4224 Human foe
C 227	36	76.6	5845	6	ABQ60990	Abq60990 2 POMH2 p	C 300	35	74.5	473	4	ABA66788	Abas66788 Human foe
C 228	36	76.6	5881	10	ADD01241	Add01241 Human nuc	C 301	35	74.5	473	6	ABS14903	Abel14903 Human gen
C 229	36	76.6	6040	12	ADP03042	Adp03042 Human hou	C 302	35	74.5	473	6	ABS02399	Abso2399 Human gen
C 230	36	76.6	6040	13	ADS88540	Ads88540 Human hou	C 303	35	74.5	499	13	ACF81577	Acf81577 Human SIR
C 231	36	76.6	6040	13	ADU60173	Adu60173 Housekeep	C 304	35	74.5	501	8	ABX50376	Abx50376 Bovine ES
C 232	36	76.6	6111	10	ADG32695	Adg32695 Human DNA	C 305	35	74.5	507	4	ABA60120	Abas60120 Human foe
C 233	36	76.6	6111	12	ADI24470	Adi24470 Human mod	C 306	35	74.5	507	4	AAI39999	Aai39999 Probe #69
C 234	36	76.6	6111	14	ADX06466	Adx06466 Cyclin-de	C 307	35	74.5	507	4	ABA28478	Abas28478 Probe #69
C 235	36	76.6	6888	3	AAA70114	Aaa70114 Pysmodiu	C 308	35	74.5	507	4	AAK34276	Aak34276 Human bon
C 236	36	76.6	7059	2	AAQ44750	Aaq44750 Soybean l	C 309	35	74.5	507	4	AAK08396	Aak08396 Human bra
C 237	36	76.6	9327	10	ADC87694	Adc87694 Human mam	C 310	35	74.5	507	4	ABQ34067	Abq34067 Human liv
C 238	36	76.6	9486	5	AAAS90982	Aaas90982 DNA encod	C 311	35	74.5	507	6	ABS09003	Abso9003 Human gen
C 239	36	76.6	9486	5	AAAS6084	Aaas6084 Novel DNA	C 312	35	74.5	518	13	ADL13521	Adl13521 Plant ful
C 240	36	76.6	9486	10	ADBE09888	Adbe09888 Novel DNA	C 313	35	74.5	602	6	ABK16233	Abk16233 Human lun
C 241	36	76.6	10196	6	AAI64255	Aai64255 Human SIC	C 314	35	74.5	602	10	ADB95496	Adb95496 Human lun
C 242	36	76.6	10302	6	AAI72720	Aai72720 BCW2 cDNA	C 315	35	74.5	603	10	ACD92677	Acd92677 Human col
C 243	36	76.6	12175	6	ABK13581	Abk13581 Ryegrass	C 316	35	74.5	603	14	ABE65850	Aebe65850 Rice geno
C 244	36	76.6	14991	4	AAAS33441	Aaas33441 DNA encod	C 317	35	74.5	610	6	ABQ66295	Abq66295 Arabidops
C 245	36	76.6	16509	6	ABL33321	Ab133321 Human imm	C 318	35	74.5	631	14	ADM84685	Adm84685 MAP3K9 ma
C 246	36	76.6	17700	4	AAK68945	Aak68945 Human imm	C 319	35	74.5	775	3	AAO20301	Aao20301 Human col
C 247	36	76.6	23213	6	ABV78027	Abv78027 Hypoxia-r	C 320	35	74.5	785	3	AAAC34315	Aaac34315 Arabidops
C 248	36	76.6	24619	13	ADP24619	Adp24619 PRO polyp	C 321	35	74.5	863	11	ACN82903	Acn82903 Breast ca
C 249	36	76.6	25356	12	ADL13021	Adl13021 Human ste	C 322	35	74.5	873	14	ADZ70641	Adz70641 Human cDN
C 250	36	76.6	39651	4	ABL18856	Ab118856 Drosophill	C 323	35	74.5	937	10	ADB81670	Adb81670 Human ova
C 251	36	76.6	51664	11	ACNA4432	Acn44432 Mouse gen	C 324	35	74.5	1002	8	ACA30235	Aca30235 Prokaryot
C 252	36	76.6	68108	14	AEBS39162	Aeb39162 L. pneumo	C 325	35	74.5	1008	3	ACA30235	Aca30235 Arabidops
C 253	36	76.6	85272	12	ADQ97064	Adq97064 Human can	C 326	35	74.5	1189	2	AAV41452	Aav41452 Nucleotid
C 254	36	76.6	85873	10	ADH10008	Adh10008 Human chr	C 327	35	74.5	1189	5	AAFP98424	Aaf98424 Human cDN
C 255	36	76.6	103464	13	ABD33278	Abd33278 Murine ca	C 328	35	74.5	1245	10	ADH82305	Adh82305 Enterococ
C 256	36	76.6	109906	6	ABK94411	Abk94411 DNA encod	C 329	35	74.5	1264	11	ADU51154	Adu51154 Murine sp
C 257	36	76.6	109906	12	ADL08112	Adl08112 Human gen	C 330	35	74.5	1264	13	ADS82758	Ads82758 Human lym
C 258	36	76.6	110000	2	AAK91990_01	Continuation (2 of	C 331	35	74.5	1276	13	ADU51158	Adu51158 Murine sp
C 259	36	76.6	110000	2	AAK91990_02	Continuation (2 of	C 332	35	74.5	1299	8	ACC46309	Acc46309 Human dlt
C 260	36	76.6	110000	8	AAK53224_3	Continuation (3 of	C 333	35	74.5	1407	5	AAAS90722	Aas90722 DNA encod
C 261	36	76.6	110000	11	ADP77343_12	Continuation (13 of	C 334	35	74.5	1434	5	AAI58607	Aai58607 Human pol
C 262	36	76.6	110000	11	ACNA44150_3	Continuation (4 of	C 335	35	74.5	1434	5	ADQ98825	Adq98825 DNA encod
C 263	36	76.6	110000	12	ADQ97960_0	Adq97960 Human can	C 336	35	74.5	1434	5	ADQ98825	Adq98825 Novel hum
C 264	36	76.6	110000	14	AEA61095_0	Aea61095 Human LOC	C 337	35	74.5	1434	9	ADB48585	Adb48585 Novel hum
C 265	36	76.6	110000	14	AEBS35724_0	Aeb35724 L. pneumo	C 338	35	74.5	1517	3	AAAC45615	Aac45615 Arabidops
C 266	36	76.6	110000	14	AEBS39175_20	Continuation (21 o	C 339	35	74.5	1517	6	ABN81545	Abn81545 Olfactory
C 267	36	76.6	110000	14	AEBA42401_19	Continuation (20 o	C 340	35	74.5	1519	3	AAAC32925	Aac32925 Arabidops
C 268	36	76.6	110000	14	AEBA4237_06	Continuation (7 of	C 341	35	74.5	1688	4	AAAF4673	Aaf4673 Novel pro
C 269	36	76.6	113306	10	ADC86554_0	Adc86554 Human GPC	C 342	35	74.5	1688	12	ADI23371	Adi23371 Mouse MAR
C 270	36	76.6	149480	6	ABL61947	Ab161947 Colon ade	C 343	35	74.5	1793	12	ADO35946	Ado35946 Novel mou
C 271	36	76.6	149480	6	ABL68365	Ab168365 Kidney ca	C 344	35	74.5	1825	4	ABL24410	Ab124410 Drosophill
C 272	36	76.6	149480	6	ABL61948	Ab161948 Colon ade	C 345	35	74.5	1850	8	ADA71405	Ada71405 Rice gene
C 273	36	76.6	158417	13	ADS36461	Ads36461 Human aut	C 346	35	74.5	1901	13	ADX47886	Adx47886 Plant ful
C 274	36	76.6	16043	12	ADL08127	Adl08127 Human gen	C 347	35	74.5	1962	2	AAV08824	Aav08824 Gene No.
C 275	36	76.6	170170	10	ADL13643	Adl13643 Orcearth	C 348	35	74.5	1962	13	ADX53270	Adx53270 Plant ful
C 276	36	76.6	199994	14	ABEA07496	Aea07496 CTC-200D1	C 349	35	74.5	1969	6	ABZ17524	Abz17524 Arabidops
C 277	36	76.6	260160	12	ADQ20017	Adq20017 Human sof	C 350	35	74.5	1969	6	ABZ16045	Abz16045 Arabidops
C 278	36	76.6	349980	5	AAH41225	Aah41225 Pyrococu	C 351	35	74.5	1979	4	ABL05085	Ab105085 Drosophill
C 279	35.5	75.5	255	5	ABN78420	Abn78420 Human str	C 352	35	74.5	2004	7	ADZ74746	Adz74746 Arabidops
C 280	35.5	75.5	2667	11	ADW21887	Adw21887 Rat hepat	C 353	35	74.5	2021	10	ADC85662	Adc85662 Human GPC
C 281	35	74.5	217	6	ABN64165	Abn64165 Human can	C 354	35	74.5	2021	10	ADC85662	Adc85662 Human GPC
C 282	35	74.5	278	14	ACL57501	Ac157501 Human col	C 355	35	74.5	2024	4	AAH14856	Aah14856 Human cDN
C 283	35	74.5	346	4	ABA72660	Abas72660 Human foe	C 356	35	74.5	2027	10	ADC85660	Adc85660 Human GPC
C 284	35	74.5	346	4	AAI53079	Aai53079 Probe #21	C 357	35	74.5	2056	5	ABA83011	Abas83011 Human tra
C 285	35	74.5	346	4	ABA38356	Abas38356 Probe #16	C 358	35	74.5	2056	12	ADL12684	Adl12684 Human ste
C 286	35	74.5	346	4	AAK47243	Aak47243 Human bon	C 359	35	74.5	2057	10	ADP42451	Adp42451 Human pp7
C 287	35	74.5	346	4	AAK21089	Aak21089 Human bra	C 360	35	74.5	2140	10	ADF81507	Adf81507 Leukaemia
C 288	35	74.5	346	4	ABSA46996	Abas46996 Human liv	C 361	35	74.5	2140	13	ADR25928	Adr25928 Breast ca
C 289	35	74.5	346	6	ABS21436	Abes21436 Human gen	C 362	35	74.5	2140	14	ADZ49667	Adz49667 Inseulin s
C 290	35	74.5	400	4	AAI21637	Aai21637 Human bre	C 363	35	74.5	2317	4	AAI60393	Aai60393 Human pol
C 291	35	74.5	414	8	ABX40997	Abx40997 Bovine ES	C 364	35	74.5	2473	11	ACN88849	Acn88849 Breast ca
C 292	35	74.5	438	5	ABV14246	Abv14246 Human pro	C 365	35	74.5	2573	4	ABL27646	Ab127646 Drosophill
C 293	35	74.5	456	5	ABV35339	Abv35339 Human pro	C 366	35	74.5	2689	8	ACN03993	Acn03993 cDNA down
C 294	35	74.5	465	13	AEA01320	Aea01320 Oryza spe	C 367	35	74.5	2689	9	ACH04173	Ach04173 Human cDN
C 295	35	74.5	466	6	ABK30806	Abk30806 Plant dwa	C 368	35	74.5	2771	14	ADZ36889	Adz36889 C4H promo
C 296	35	74.5	470	4	AAI12766	Aai12766 Human bre	C 369	35	74.5	3015	13	ADX29049	Adx29049 Plant ful

370	35	74.5	3087	11	ACL28826	Adz28826	Rice abio	443	35	74.5	186510	10	ADB24797	Adz24797	Human ena
371	35	74.5	3552	14	ADZ36899	Adz36899	C4H promo	C 444	35	74.5	197140	14	AEA17302	Adz17302	Human GNA
372	35	74.5	4284	4	ABL05084	Adz05084	Drosophila	445	35	74.5	201239	8	ACA64924	Adz64924	Human PLZ
373	35	74.5	4431	4	ABL29958	Adz29958	Drosophila	446	35	74.5	208700	13	ABD32688	Adz32688	Human can
374	35	74.5	4940	2	AAV21451	P. falcip	447	35	74.5	240102	13	ABD32546	Adz32546	Mouse can	
375	35	74.5	5289	8	ABT17821	Adz17821	Aspergill	C 448	35	74.5	241748	14	ADZ13116	Adz13116	Murine ca
376	35	74.5	5289	8	ABT19635	Adz19635	Aspergill	449	35	74.5	260803	13	ABD32730	Adz32730	Human can
377	35	74.5	5340	4	ABL22772	Adz22772	Drosophila	C 450	35	74.5	295096	11	ACN44068	Adz44068	Mouse gen
378	35	74.5	5432	2	AAV10236	Adz10236	Arabidops	451	35	74.5	337022	12	ADQ59416	Adz59416	Human can
379	35	74.5	5942	6	ABK33945	Adz33945	Human DNA	452	35	74.5	338780	14	ADZ13691	Adz13691	Human can
380	35	74.5	5942	6	ABK31223	Adz31223	Signal tr	C 453	34	72.3	25	9	ACI64559	Adz64559	Human mic
381	35	74.5	5942	6	ABK170534	Adz170534	Chemical tr	C 454	34	72.3	48	4	ABK09220	Adz09220	Human CD2
382	35	74.5	5942	6	AA611135	Adz611135	Human gen	C 455	34	72.3	48	4	ABK07037	Adz07037	Human NOG
383	35	74.5	5942	8	ABZ10002	Adz10002	Haematopo	C 456	34	72.3	48	5	ADV07095	Adz07095	Human BAC
384	35	74.5	5942	8	ABZ10088	Adz10088	Haematopo	C 457	34	72.3	48	11	ADL53188	Adz53188	Human NOG
385	35	74.5	5942	8	ADA20351	Adz20351	Prostate	C 458	34	72.3	48	11	ADL53220	Adz53220	Human NOG
386	35	74.5	5942	8	ADA84158	Adz84158	Human ren	459	34	72.3	51	4	AAI28858	Adz28858	Human SNP
387	35	74.5	5942	10	ADZ84068	Adz84068	Human lym	460	34	72.3	51	4	AAI28857	Adz28857	Human SNP
388	35	74.5	6285	2	AAI13352	Adz13352	Enterococ	C 461	34	72.3	228	4	AAK77241	Adz77241	Human imm
389	35	74.5	6285	6	ABS99147	Adz99147	Enterococ	C 462	34	72.3	314	4	AAK60239	Adz60239	Human imm
390	35	74.5	6682	13	ADR84830	Adz84830	Aspergill	463	34	72.3	372	12	ADP92113	Adz92113	Cotton ex
391	35	74.5	7093	4	AAH24955	Adz24955	Nucleotid	464	34	72.3	386	2	AAV87326	Adz87326	EST clone
392	35	74.5	7093	6	AAD22525	Adz22525	Mouse Tie	465	34	72.3	388	3	AAH30214	Adz30214	Human col
393	35	74.5	8132	13	ADR84205	Adz84205	Aspergill	466	34	72.3	404	8	ABZ52984	Adz52984	Aspergill
394	35	74.5	9646	6	ABZ22017	Adz22017	Human leu	C 467	34	72.3	432	6	ABL79967	Adz79967	Human ova
395	35	74.5	12682	13	ADR84243	Adz84243	Aspergill	468	34	72.3	453	5	ABV58149	Adz58149	Human pro
396	35	74.5	13074	14	ABZ21818	Adz21818	E. coli O	469	34	72.3	457	3	AAV31434	Adz31434	Human sec
397	35	74.5	14304	6	ABK69840	Adz69840	Human sec	470	34	72.3	480	4	AAI11757	Adz11757	Probe #16
398	35	74.5	15531	4	AAAS31521	Adz31521	Human DNA	471	34	72.3	480	4	ABA53448	Adz53448	Human foe
399	35	74.5	15531	5	ABA17065	Adz17065	Human ner	472	34	72.3	480	4	AAI33067	Adz33067	Probe #17
400	35	74.5	15531	6	ABQ66845	Adz66845	Human pol	473	34	72.3	480	4	ABA43035	Adz43035	Human bre
401	35	74.5	15531	10	ADCL1132	Adz1132	Human DNA	474	34	72.3	480	4	ABA23221	Adz23221	Probe #16
402	35	74.5	15962	6	ABA01445	Adz01445	Streptoco	475	34	72.3	480	4	AAK01719	Adz01719	Human bra
403	35	74.5	19738	6	ABA01436	Adz01436	Streptoco	476	34	72.3	480	4	AAK01719	Adz01719	Human bra
404	35	74.5	20001	13	ADZ77150	Adz77150	Type II d	477	34	72.3	480	4	ABS26754	Adz26754	Human liv
405	35	74.5	20905	4	ABA07327	Adz07327	Human pan	478	34	72.3	480	5	AAI01686	Adz01686	Probe #16
406	35	74.5	20905	4	AAK90486	Adz90486	Human dig	479	34	72.3	480	6	ABS01722	Adz01722	Human gen
407	35	74.5	20905	4	AAK87167	Adz87167	Human imm	C 480	34	72.3	490	3	AAAF08267	Adz08267	Fusarium
408	35	74.5	26197	4	ABK43078	Adz43078	Genomic s	481	34	72.3	490	4	AAI12986	Adz12986	Probe #29
409	35	74.5	26197	9	ADB61234	Adz61234	Connectiv	482	34	72.3	490	4	ABA54686	Adz54686	Human foe
410	35	74.5	26210	4	ABK43079	Adz43079	Genomic s	483	34	72.3	490	4	AAI34343	Adz34343	Probe #30
411	35	74.5	26210	9	ADB61235	Adz61235	Connectiv	484	34	72.3	490	4	ABA44238	Adz44238	Human bra
412	35	74.5	27433	11	ACN44648	Adz44648	Mouse gen	485	34	72.3	490	4	ABA24469	Adz24469	Probe #29
413	35	74.5	30277	11	ACN44648	Adz44648	Mouse gen	486	34	72.3	490	4	AAK28418	Adz28418	Human bon
414	35	74.5	31050	12	ADQ97078	Adz97078	Human can	487	34	72.3	490	4	AAK02973	Adz02973	Human bra
415	35	74.5	35183	13	ABD33355	Adz33355	Murine ca	488	34	72.3	490	4	ABS28015	Adz28015	Human liv
416	35	74.5	39776	10	AAI51353	Adz51353	Human sec	489	34	72.3	490	5	AAI02903	Adz02903	Probe #28
417	35	74.5	43284	14	ADZ70415	Adz70415	Human cdn	490	34	72.3	490	6	ABS02926	Adz02926	Human gen
418	35	74.5	43599	6	ABK84242	Adz84242	Human cdn	491	34	72.3	490	9	ACH19018	Adz19018	Human edu
419	35	74.5	44456	4	ABL09962	Adz09962	Drosophila	C 492	34	72.3	490	13	ADU52308	Adz52308	Fusarium
420	35	74.5	58329	11	ACN43902	Adz43902	Human gen	C 493	34	72.3	490	14	ADZ90311	Adz90311	Fusarium
421	35	74.5	64316	12	ADQ97357	Adz97357	Human can	494	34	72.3	493	8	ABZ54636	Adz54636	Aspergill
422	35	74.5	82660	11	ACN45192	Adz45192	Mouse gen	C 495	34	72.3	544	6	ABN60734	Adz60734	Human can
423	35	74.5	83493	14	ADZ13310	Adz13310	Murine ca	496	34	72.3	545	6	ABQ35409	Adz35409	Oligonuc
424	35	74.5	86804	12	ADQ97700	Adz97700	Mouse can	C 497	34	72.3	545	6	ABQ35408	Adz35408	Oligonuc
425	35	74.5	89500	12	ADQ56275	Adz56275	Human pre	C 498	34	72.3	563	14	ACL58953	Adz58953	Human col
426	35	74.5	89700	14	ADZ80725	Adz80725	Human pre	499	34	72.3	568	6	ABQ57144	Adz57144	Human col
427	35	74.5	99046	13	ABD33291	Adz33291	Human can	500	34	72.3	568	6	ABZ14981	Adz14981	Arabidops
428	35	74.5	101513	14	ADZ13341	Adz13341	Human can	501	34	72.3	615	6	ABK75632	Adz75632	Bacillus
429	35	74.5	110000	5	AAI61373	Adz61373	Continuation (5 of	502	34	72.3	622	12	ADN13820	Adz13820	Human pro
430	35	74.5	110000	12	ADN46845	Adz46845	Continuation (15 of	503	34	72.3	644	6	ABN89038	Adz89038	Human pro
431	35	74.5	110000	12	ADN47591	Adz47591	Continuation (7 of	C 504	34	72.3	678	6	ABN90915	Adz90915	Staphyloc
432	35	74.5	110000	12	ADN46123	Adz46123	Continuation (15 of	505	34	72.3	685	8	ABZ51427	Adz51427	Aspergill
433	35	74.5	110000	12	ADN47209	Adz47209	Continuation (7 of	C 506	34	72.3	687	3	AAF22107	Adz22107	Arabidops
434	35	74.5	110000	12	ADN46464	Adz46464	Continuation (15 of	C 507	34	72.3	724	4	AAH05855	Adz05855	Human cdn
435	35	74.5	110000	12	ADN47960	Adz47960	Continuation (7 of	508	34	72.3	756	4	AAK73431	Adz73431	Human imm
436	35	74.5	110000	12	ADQ97331	Adz97331	Continuation (2 of	509	34	72.3	756	6	ABQ35229	Adz35229	Oligonuc
437	35	74.5	110000	12	ADQ97331	Adz97331	Continuation (3 of	C 510	34	72.3	756	6	ABQ35228	Adz35228	Oligonuc
438	35	74.5	110000	4	ADZ12821	Adz12821	Continuation (2 of	511	34	72.3	771	13	ADU26310	Adz26310	meho gene
439	35	74.5	112190	4	AAH44801	Adz44801	Human GPC	512	34	72.3	781	3	AAK95506	Adz95506	Human sec
440	35	74.5	128963	12	ADQ97110	Adz97110	Human can	C 513	34	72.3	866	4	AAK88356	Adz88356	H. tuberc
441	35	74.5	129017	12	ADP84158	Adz84158	Human AST	C 514	34	72.3	891	13	ADK63821	Adz63821	Plant ful
442	35	74.5	163319	3	AAF22306	Adz22306	Arabidops	515	34	72.3	895	13	ADR63788	Adz63788	Cotton cd

C 516	34	72.3	945	6	ABQ69622	Abq69622 Listeria	589	34	72.3	2564	12	ADQ64910	Adq64910 Novel hum
C 517	34	72.3	945	6	ABQ69514	Abq69514 Listeria	C 590	34	72.3	2633	6	ABQ61076	Abq61076 RIKEN 120
C 518	34	72.3	945	6	ABQ67734	Abq67734 Listeria	C 591	34	72.3	2635	13	ADR37736	Adr37736 Human chr
C 519	34	72.3	963	6	ABQ21446	Abq21446 Oligonucle	C 592	34	72.3	2635	13	ADR37734	Adr37734 Human chr
C 520	34	72.3	963	6	ABQ21447	Abq21447 Oligonucle	C 593	34	72.3	2642	4	ABL13712	Ab113712 Drosophil
C 521	34	72.3	1025	4	AAS25995	Aas25995 Human cDN	C 594	34	72.3	2644	12	ADK51975	Adk51975 Human ato
C 522	34	72.3	1025	6	ABX73336	Abx73336 Human nov	C 595	34	72.3	2644	12	ADN04839	Adn04839 Antipsori
C 523	34	72.3	1164	8	ABX70288	Abx70288 Listeria	C 596	34	72.3	2644	14	ADY117425	Ady117425 DNA encod
C 524	34	72.3	1188	8	ACA40081	Aca40081 Prokaryot	C 597	34	72.3	2688	5	ABV27183	Abv27183 Human pro
C 525	34	72.3	1252	2	ACA56617	Aca56617 Equine he	C 598	34	72.3	2688	5	ABV21364	Abv21364 Human pro
C 526	34	72.3	1252	2	AAT00532	Aat00532 Equine he	C 599	34	72.3	2696	11	ACN91875	Acn91875 Breast ca
C 527	34	72.3	1278	6	ABQ79279	Abq79279 Polypepti	C 600	34	72.3	2720	3	AAZ94148	Aaz94148 Rat TPL-2
C 528	34	72.3	1306	2	AAQ27827	Aaq27827 Bovine TP	C 601	34	72.3	2720	5	AAZ17450	Aaz17450 Rat Tpl2
C 529	34	72.3	1307	2	AAK61083	Aak61083 Tomato (1	C 602	34	72.3	2720	12	ADI119734	Adi119734 Rat tumou
C 530	34	72.3	1365	4	ABL21899	Ab121899 Drosophil	C 603	34	72.3	2730	3	AAZ64225	Aaz64225 Human pro
C 531	34	72.3	1384	3	AAZ64227	Aaz64227 Human pol	C 604	34	72.3	2768	10	ADK53831	Adk53831 Human pro
C 532	34	72.3	1421	6	ABX97070	Abx97070 Human NOV	C 605	34	72.3	2769	13	ACF87490	Acf87490 Human STR
C 533	34	72.3	1482	11	ACN44145	Acn44145 Mouse mRN	C 606	34	72.3	2958	4	AAK52250	Aak52250 Human pol
C 534	34	72.3	1539	8	ACA31831	Aca31831 Prokaryot	C 607	34	72.3	3001	3	AAH51714	Aah51714 Chromosom
C 535	34	72.3	1556	12	ADQ17078	Adq17078 Porcine F	C 608	34	72.3	3002	4	ABL09763	Ab109763 Drosophil
C 536	34	72.3	1590	8	ACA26220	Aca26220 Prokaryot	C 609	34	72.3	3008	11	ACN90791	Acn90791 Breast ca
C 537	34	72.3	1608	3	AAZ34484	Aaz34484 Arabidops	C 610	34	72.3	3064	3	AAF15575	Aaf15575 Human pro
C 538	34	72.3	1710	2	AAT39797	Aat39797 Human clo	C 611	34	72.3	3090	8	ADA70829	Ada70829 Rice gene
C 539	34	72.3	1740	14	ADV42575	Adv42575 Human psy	C 612	34	72.3	3096	12	ADL82936	Adl82936 Human PRO
C 540	34	72.3	1781	10	ADD12695	Add12695 Human cDN	C 613	34	72.3	3096	13	ADP54539	Adp54539 Human PRO
C 541	34	72.3	1826	13	ADT44238	Adt44238 Bacterial	C 614	34	72.3	3096	14	ADY19845	Ady19845 DNA encod
C 542	34	72.3	1850	4	ABA44272	Ab444272 Human bre	C 615	34	72.3	3143	13	ADW50754	Adw50754 cDNA sequ
C 543	34	72.3	1850	4	AAK03006	Aak03006 Human bra	C 616	34	72.3	3144	4	ABN87474	Abn87474 Human zin
C 544	34	72.3	1877	13	ADX61860	Adx61860 Plant ful	C 617	34	72.3	3321	13	ADQ07314	Adq07314 Full leng
C 545	34	72.3	1909	13	ADX62458	Adx62458 Plant ful	C 618	34	72.3	3420	8	ABX34662	Abx34662 Human mdd
C 546	34	72.3	1986	6	ABK74552	Abk74552 Bacillus	C 619	34	72.3	3424	4	ABL21898	Ab121898 Drosophil
C 547	34	72.3	2000	7	ADZ75333	Adz75333 Rice prom	C 620	34	72.3	3442	4	ABL30096	Ab130096 Drosophil
C 548	34	72.3	2000	8	ADA72737	Ada72737 Rice gene	C 621	34	72.3	3467	3	AAZ91907	Aaz91907 Human pro
C 549	34	72.3	2036	4	ABL08287	Ab108287 Drosophil	C 622	34	72.3	3467	3	AAZ919132	Aaz919132 LAR tyros
C 550	34	72.3	2051	6	ABQ79280	Abq79280 Polypepti	C 623	34	72.3	3693	13	ADQ04090	Adq04090 Potato st
C 551	34	72.3	2055	14	ADZ44598	Adz44598 ReJB DNA	C 624	34	72.3	3719	12	ADO35658	Ado35658 Novel mou
C 552	34	72.3	2163	13	ADX46124	Adx46124 Plant ful	C 625	34	72.3	3769	12	ADH56320	Adh56320 Human S30
C 553	34	72.3	2198	10	ADC08366	Adc08366 Rice DNA	C 626	34	72.3	3858	2	AAZ20531	Aaz20531 Polynucle
C 554	34	72.3	2203	10	ADDA4371	Ada4371 Human gen	C 627	34	72.3	3867	3	AAZ93224	Aaz93224 Fatty aci
C 555	34	72.3	2203	10	ADDA47209	Ada47209 Human gen	C 628	34	72.3	4012	4	ABL14202	Ab114202 Drosophil
C 556	34	72.3	2205	5	ABA82994	Ab82994 Human tra	C 629	34	72.3	4122	10	ADC99109	Adc99109 Human RFP
C 557	34	72.3	2232	4	AAH14261	Aah14261 Human cDN	C 630	34	72.3	4152	8	ACD13376	Ac13376 Human DNA
C 558	34	72.3	2232	8	ABX76198	Abx76198 Lung canc	C 631	34	72.3	4152	12	ADMI1398	Adm11398 Human CDC
C 559	34	72.3	2232	13	ADR25886	Adr25886 Breast ca	C 632	34	72.3	4152	12	ADMO1240	Ado11240 Human CDC
C 560	34	72.3	2238	13	ADT43452	Adt43452 Bacterial	C 633	34	72.3	4167	12	ADQ26019	Adq26019 Potato st
C 561	34	72.3	2284	12	ADQ86959	Adq86959 Human tum	C 634	34	72.3	4167	12	ADQ76437	Adq76437 Nucleotid
C 562	34	72.3	2302	2	AAK02971	Aak02971 Human II-	C 635	34	72.3	4262	13	ADP54561	Adp54561 Human PRO
C 563	34	72.3	2314	6	ABL67654	Ab167654 Oesophagu	C 636	34	72.3	4262	14	ADY15519	Ady15519 DNA encod
C 564	34	72.3	2314	6	ABV94428	Abv94428 Breast ca	C 637	34	72.3	4386	4	ABL08286	Ab108286 Drosophil
C 565	34	72.3	2314	6	ABK84095	Abk84095 Human cDN	C 638	34	72.3	4392	10	ADDE60010	Adde60010 Human gen
C 566	34	72.3	2314	8	ABX10955	Abx10955 cDNA sequ	C 639	34	72.3	4392	10	ADD45761	Add45761 Human gen
C 567	34	72.3	2314	13	ADR25432	Adr25432 Breast ca	C 640	34	72.3	4409	4	ABL05825	Ab105825 Drosophil
C 568	34	72.3	2324	10	ABV99861	Abv99861 Human 121	C 641	34	72.3	4428	4	AAQ06575	Aaq06575 Bovine al
C 569	34	72.3	2325	13	ADK62896	Adk62896 Plant ful	C 642	34	72.3	4600	13	ADR06902	Adr06902 Full leng
C 570	34	72.3	2328	4	AAK53234	Aak53234 Human pol	C 643	34	72.3	4836	12	ADQ63853	Adq63853 Novel hum
C 571	34	72.3	2328	4	AAK53031	Aak53031 Human dia	C 644	34	72.3	4992	5	ABV29871	Abv29871 Human pro
C 572	34	72.3	2427	13	ADX27540	Adx27540 Plant ful	C 645	34	72.3	4992	5	ABV23988	Abv23988 Human pro
C 573	34	72.3	2473	10	ABV99868	Abv99868 Human 121	C 646	34	72.3	5755	4	ABL14212	Ab114212 Drosophil
C 574	34	72.3	2473	10	ABV99865	Abv99865 Human 121	C 647	34	72.3	5725	14	ADU77968	Adu77968 Human apo
C 575	34	72.3	2473	10	ABV99860	Abv99860 Human 121	C 648	34	72.3	6061	4	AAK45334	Aak45334 Chemical
C 576	34	72.3	2473	10	ABV99867	Abv99867 Human 121	C 649	34	72.3	6220	14	ADY26915	Ady26915 Human vim
C 577	34	72.3	2473	10	ABV99863	Abv99863 Human 121	C 650	34	72.3	6254	6	ABL33620	Ab133620 Human imm
C 578	34	72.3	2473	10	ABV99862	Abv99862 Human 121	C 651	34	72.3	6259	2	AAK86366	Aak86366 SN22 prot
C 579	34	72.3	2473	10	ABV99864	Abv99864 Human 121	C 652	34	72.3	6424	4	ABL09762	Ab109762 Drosophil
C 580	34	72.3	2473	10	ABV99866	Abv99866 Human 121	C 653	34	72.3	6488	4	AAK87050	Aak87050 Human imm
C 581	34	72.3	2496	8	ABX72225	Abx72225 Human NOV	C 654	34	72.3	6488	4	AAK65490	Aak65490 Human imm
C 582	34	72.3	2496	9	ADA38099	Ada38099 Map3K8 DN	C 655	34	72.3	7007	12	ADQ85067	Adq85067 Human tum
C 583	34	72.3	2496	11	ADM10600	Adm10600 Human Map	C 656	34	72.3	7142	4	ABL17798	Ab117798 Drosophil
C 584	34	72.3	2507	9	ADA02838	Ada02838 Mouse Map	C 657	34	72.3	7266	13	ADR84508	Adr84508 Aspergill
C 585	34	72.3	2507	10	ADB72576	Adb72576 Mouse Map	C 658	34	72.3	7266	13	ADR84385	Adr84385 Aspergill
C 586	34	72.3	2507	10	ADC85317	Adc85317 Mouse Map	C 659	34	72.3	7266	13	ADR84386	Adr84386 Aspergill
C 587	34	72.3	2507	12	ADM74433	Adm74433 Murine ca	C 660	34	72.3	7471	4	ABL05824	Ab105824 Drosophil
C 588	34	72.3	2558	4	ABL23998	Ab123998 Drosophil	C 661	34	72.3	7702	3	AAZ91908	Aaz91908 Human pro

C 662	34	72.3	7702	3	AAZ59133	Aaz59133 LAR tyros	735	34	72.3	110000	6	ABQ69245_17	Continuation (18 o	
C 663	34	72.3	7702	3	AAA88739	Aaa88739 Human pro	C 736	34	72.3	110000	6	ABQ67195_0	ABQ67195 Listeria	
C 664	34	72.3	7702	10	ADD18741	Add18741 Human dis	737	34	72.3	110000	6	ABA03041_15	Continuation (16 o	
C 665	34	72.3	7702	10	ADK61221	Adk61221 Ovarian c	738	34	72.3	110000	8	AAZ53223_2	Continuation (3 of	
C 666	34	72.3	7702	11	ADI32010	Adi32010 Human cdn	739	34	72.3	110000	8	AAZ53223_3	Continuation (4 of	
C 667	34	72.3	7702	12	ADP18653	Adp18653 Human TAT	740	34	72.3	110000	10	ADF77343_06	Continuation (7 of	
C 668	34	72.3	7702	13	ADJ33475	Adj33475 Human leu	C 741	34	72.3	110000	11	ACN44582_0	ACN44582 Human gen	
C 669	34	72.3	7702	13	ADJ33475	Adj33475 Human leu	C 742	34	72.3	110000	13	ABD32966_04	Continuation (5 of	
C 670	34	72.3	7705	4	AAH98405	Aah98405 Human EST	C 743	34	72.3	110000	13	ABD32966_09	Continuation (10 o	
C 671	34	72.3	7705	4	AAH98405	Aah98405 Human EST	C 744	34	72.3	110000	13	ABD32535_0	ABD32535 Human can	
C 672	34	72.3	7718	13	ADV35116	Adv35116 Human cdn	745	34	72.3	110000	13	ABD32806_3	Continuation (4 of	
C 673	34	72.3	7724	13	ADJ33489	Adj33489 Human LAR	746	34	72.3	110000	13	ADV81204_11	Continuation (12 o	
C 674	34	72.3	7741	4	AAZ52448	Aaz52448 Human cdn	C 747	34	72.3	110000	14	ADZ13035_0	ADZ13035 Human can	
C 675	34	72.3	7945	5	ABV27897	Abv27897 Human pro	C 748	34	72.3	110000	14	ABE39175_11	Continuation (12 o	
C 676	34	72.3	8232	11	ACN89826	Acn89826 Breast ca	C 749	34	72.3	110000	14	ABE39175_22	Continuation (23 o	
C 677	34	72.3	8738	4	ABL11376	Ab111376 Drosophil	C 750	34	72.3	110000	14	ABE42401_11	Continuation (12 o	
C 678	34	72.3	8920	4	ABL17648	Ab117648 Drosophil	C 751	34	72.3	110000	14	ABE42401_19	Continuation (20 o	
C 679	34	72.3	10102	4	AAK83072	Aak83072 Human imm	C 752	34	72.3	110000	14	ABE42401_21	Continuation (22 o	
C 680	34	72.3	10497	2	AAK03048	Aak03048 Human IL-	C 753	34	72.3	110000	14	ABE42737_08	Continuation (9 of	
C 681	34	72.3	10906	2	AAV31249	Aav31249 E. coli J	C 754	34	72.3	115223	13	ABD33568	ABD33568 Murine ca	
C 682	34	72.3	14752	2	AAK02056	Aak02056 Borrelia	C 755	34	72.3	127197	5	ABT161370	ABT161370 Soybean s	
C 683	34	72.3	19480	4	AAK80384	Aak80384 Human imm	C 756	34	72.3	140167	6	ABT10146	ABT10146 Human bre	
C 684	34	72.3	19481	4	AAK80383	Aak80383 Human imm	757	34	72.3	142299	10	ADD50651	Add50651 BAC sequ	
C 685	34	72.3	19734	6	ABL33933	Ab133933 Human imm	758	34	72.3	142299	14	ADV77909	Adv77909 Human BAC	
C 686	34	72.3	20566	4	AAK87596	Aak87596 Human imm	C 759	34	72.3	144411	12	ADP74214	Adp74214 Equine he	
C 687	34	72.3	21587	13	ABD33127	Abd33127 Murine ca	C 760	34	72.3	144486	12	ADP74215	Adp74215 Equine he	
C 688	34	72.3	22644	11	ACN44144	Acn44144 Mouse gen	C 761	34	72.3	145444	12	ADP74213	Adp74213 Equine he	
C 689	34	72.3	23101	14	ADZ13394	Adz13394 Murine ca	C 762	34	72.3	145596	12	ADP74202	Adp74202 Equine he	
C 690	34	72.3	23451	9	ACD19160	Ac19160 E. coli 0	763	34	72.3	147300	12	ADP45593	Adp45593 Human rho	
C 691	34	72.3	23454	10	ADC01084	Adc01084 Enterohae	764	34	72.3	147700	14	ADX98570	Adx98570 Human gua	
C 692	34	72.3	25003	5	ABA19679	Ab19679 Human ner	765	34	72.3	163350	6	AD46127	Ad46127 Human tum	
C 693	34	72.3	28871	13	ADT05539	Adt05539 Haemophil	766	34	72.3	189430	14	ABE35718	AbE35718 L. pneumo	
C 694	34	72.3	29209	13	ADV87715	Adv87715 Streptoco	767	34	72.3	198849	14	ADZ13007	Adz13007 Human can	
C 695	34	72.3	29209	13	ADV87968	Adv87968 Streptoco	768	34	72.3	199377	10	ADC35071	Adc35071 Mouse gen	
C 696	34	72.3	31562	4	ABL03868	Ab103868 Drosophil	769	34	72.3	225734	12	ADQ59377	Adq59377 Human can	
C 697	34	72.3	32247	5	ABA19669	Ab19669 Human ner	770	34	72.3	225734	14	ADZ13617	Adz13617 Murine ca	
C 698	34	72.3	32934	11	ACN44722	Acn44722 Human gen	C 771	34	72.3	256157	11	ACN44650	Acn44650 Human gen	
C 699	34	72.3	34161	11	ACN44874	Acn44874 Human gen	C 772	34	72.3	256157	13	ABD33570	ABD33570 Human can	
C 700	34	72.3	34161	13	ABD33349	Abd33349 Human can	773	34	72.3	256525	11	ACN44148	Acn44148 Mouse gen	
C 701	34	72.3	41637	9	ADA02837	Ada02837 Mouse Map	C 774	34	72.3	260803	13	ABD32730	ABD32730 Human can	
C 702	34	72.3	41637	10	ADB27575	Adb27575 Mouse Map	C 775	34	72.3	263744	10	ADF08271	Adf08271 Mouse apo	
C 703	34	72.3	41637	10	ADC85316	Adc85316 Human Mef	C 776	34	72.3	263853	14	ABE39171	ABE39171 L. pneumo	
C 704	34	72.3	41637	12	ADW74432	Adw74432 Murine ca	C 777	34	72.3	276820	11	ADP75188	Adp75188 Human ADA	
C 705	34	72.3	47573	9	ADA02840	Ada02840 Human MAP	778	34	72.3	308766	13	ADT05738	Adt05738 Haemophil	
C 706	34	72.3	47573	10	ADB27578	Adb27578 Human MAP	C 779	34	72.3	340449	8	AAZ52198	AAZ52198 Human sec	
C 707	34	72.3	47573	10	ADC85319	Adc85319 Mouse Map	C 780	33.5	71.3	8048	4	AAK91437	AAK91437 Human dig	
C 708	34	72.3	47573	12	ADW74435	Adw74435 Human car	C 781	33.5	71.3	8048	5	AAZ40020	AAZ40020 Genomic s	
C 709	34	72.3	48452	4	ABL07108	Ab107108 Drosophil	C 782	33.5	71.3	8048	9	ADB32980	ADB32980 Human nov	
C 710	34	72.3	49145	13	ABD32564	Abd32564 Mouse can	C 783	33	70.2	98	2	AAQ81925	AAQ81925 Interfero	
C 711	34	72.3	50570	13	ABD32791_3	Abd32791_3 Continuation (4 of	C 784	33	70.2	113	3	AAA82095	AAA82095 N. mening	
C 712	34	72.3	60452	11	ACN45114	Acn45114 Human gen	C 785	33	70.2	114	1	AAH81098	AAH81098 DNA encod	
C 713	34	72.3	63974	14	ABE35707	AbE35707 L. pneumo	C 786	33	70.2	114	1	AAH82055	AAH82055 DNA encod	
C 714	34	72.3	75976	13	ABD33217	Abd33217 Murine ca	C 787	33	70.2	142	4	AAK85832	AAK85832 Human imm	
C 715	34	72.3	78025	14	ADZ13607	Adz13607 Murine ca	C 788	33	70.2	142	4	AAK85834	AAK85834 Human imm	
C 716	34	72.3	86131	10	ADF77178	Adf77178 KALPA gen	C 789	33	70.2	201	13	ADQ43994	Adq43994 Myocardia	
C 717	34	72.3	86765	10	ADD14752	Add14752 Human src	C 790	33	70.2	201	13	ADQ43993	Adq43993 Myocardia	
C 718	34	72.3	89014	14	ABE77360	AbE77360 Human TGF	C 791	33	70.2	221	6	ABE78871	ABE78871 E. coli C	
C 719	34	72.3	90401	12	ADQ97515	Adq97515 Human can	C 792	33	70.2	221	10	ADH80438	Adh80438 Escherich	
C 720	34	72.3	92112	13	ADG99457_3	Adg99457_3 Continuation (4 of	C 793	33	70.2	240	3	AAK32160	AAK32160 Human sec	
C 721	34	72.3	94001	13	ADJ33491	Adj33491 Human LAR	C 794	33	70.2	261	4	AAI04131	AAI04131 Human rep	
C 722	34	72.3	96594	10	ADC85476	Adc85476 Human Mef	C 795	33	70.2	261	5	AAZ40506	AAZ40506 DNA encod	
C 723	34	72.3	96595	9	ADA02996	Ada02996 Human Mef	C 796	33	70.2	261	11	ADJ09712	Adj09712 Human pro	
C 724	34	72.3	96595	12	ADW72734	Adw72734 Human Mef	C 797	33	70.2	263	3	AAZ4694	AAZ4694 Human sec	
C 725	34	72.3	96595	12	ADW74591	Adw74591 Human car	C 798	33	70.2	269	6	ABL85447	ABL85447 Human ova	
C 726	34	72.3	96988	3	AAF22290	Aaf22290 BAC conta	C 799	33	70.2	276	6	ABK77900	ABK77900 Bacillus	
C 727	34	72.3	101193	13	ABD33370	Abd33370 Human can	C 800	33	70.2	277	5	AAH80567	AAH80567 DNA encod	
C 728	34	72.3	101786	3	AAF22293	Aaf22293 BAC conta	C 801	33	70.2	282	3	AAA42661	AAA42661 Human sec	
C 729	34	72.3	10973	3	AAZ22298	Aaz22298 BAC conta	C 802	33	70.2	282	309	5	AD176020	AD176020 Human ova
C 730	34	72.3	110000	2	AAZ42063_00	AAZ42063 Haemophil	C 803	33	70.2	309	5	AD169683	AD169683 Human ova	
C 731	34	72.3	110000	2	AAZ21209_06	AAZ21209_06 Continuation (7 of	C 804	33	70.2	313	5	ABV17636	ABV17636 Human pro	
C 732	34	72.3	110000	2	AAZ01429_06	AAZ01429_06 Continuation (7 of	C 805	33	70.2	316	6	ABN20956	ABN20956 Human ORF	
C 733	34	72.3	110000	6	ABA92787_4	ABA92787_4 Continuation (5 of	C 806	33	70.2	328	6	ABN15897	ABN15897 Human ORF	
C 734	34	72.3	110000	6	ABQ69245_12	ABQ69245_12 Continuation (13 o	C 807	33	70.2	343	6	ABN19261	ABN19261 Human ORF	

808	33	70.2	344	9	ACH17462	Ach17462 Human adu	c 881	33	70.2	567	8	ACA28894	ACA28894 Prokaryot
809	33	70.2	346	3	ABQ62928	Abq62928 Mycobacte	c 882	33	70.2	572	3	AA42744	AA42744 Human sec
c 810	33	70.2	349	5	ABV15795	Abv15795 Human pro	c 883	33	70.2	573	10	ADC90828	Adc90828 E. faeciu
c 811	33	70.2	350	10	ABZ69415	Abz69415 Human CDK	c 884	33	70.2	573	13	ACN45638	Acn45638 Cotton pr
c 812	33	70.2	362	7	ADST72922	Adst72922 Human kid	c 885	33	70.2	576	14	ACL62206	ACL62206 Human col
c 813	33	70.2	362	7	ADW41776	Adw41776 cDNA elev	c 886	33	70.2	577	4	AAH34082	AAH34082 Human col
c 814	33	70.2	364	13	ADX15156	Adx15156 Plant ful	c 887	33	70.2	578	12	ACH74279	Ach74279 Human gen
c 815	33	70.2	367	9	ACH36842	Ach36842 Human end	c 888	33	70.2	610	13	ACN45805	Acn45805 Cotton pr
c 816	33	70.2	367	14	ADW05978	Adw05978 Human gen	c 889	33	70.2	610	13	ACN45805	Acn45805 Cotton pr
c 817	33	70.2	368	5	ADL41256	Adl41256 Human ova	c 890	33	70.2	615	11	ACN88350	Acn88350 Breast ca
c 818	33	70.2	370	13	ACF85151	Acf85151 Human SIR	c 891	33	70.2	617	14	ADX38735	Adx38735 Human can
c 819	33	70.2	371	4	AAL34954	Aal34954 Human mus	c 892	33	70.2	621	11	ACL31899	ACL31899 Rice abio
c 820	33	70.2	371	8	ABX57942	Abx57942 cDNA enco	c 893	33	70.2	624	13	ADX48968	Adx48968 Plant ful
c 821	33	70.2	371	12	ADJ27669	Adj27669 Human mus	c 894	33	70.2	626	3	AACT75574	Aac75574 Human ORF
c 822	33	70.2	378	6	ABN22098	Abn22098 Human ORF	c 895	33	70.2	627	5	ABV45491	Abv45491 Human pro
c 823	33	70.2	396	5	ABV35233	Abv35233 Human pro	c 896	33	70.2	628	6	ABQ66139	Abq66139 Arabidops
c 824	33	70.2	402	2	AAX91504	Aax91504 S. aureus	c 897	33	70.2	630	9	ACL24833	ACL24833 DNA clone
c 825	33	70.2	403	5	ABV14139	Abv14139 Human pro	c 898	33	70.2	645	3	AAFI1958	Aaf11958 Aspergill
c 826	33	70.2	411	13	ACF90627	Acf90627 Human SIR	c 899	33	70.2	645	13	ADU55999	Adu55999 Aspergill
c 827	33	70.2	424	5	ABV47430	Abv47430 Human pro	c 900	33	70.2	645	14	ADZ94002	Adz94002 Aspergill
c 828	33	70.2	425	2	AAX41515	Aax41515 Human sec	c 901	33	70.2	658	4	AAI86290	AAI86290 Human pol
c 829	33	70.2	425	3	ABQ63046	Abq63046 Mycobacte	c 902	33	70.2	660	8	ABZ52167	Abz52167 Aspergill
c 830	33	70.2	429	9	ACH18420	Ach18420 Human adu	c 903	33	70.2	661	13	ADQ52919	Adq52919 Novel can
c 831	33	70.2	433	9	ACH30111	Ach30111 Human tes	c 904	33	70.2	668	13	AAFI2485	Aaf12485 Aspergill
c 832	33	70.2	437	6	ABL94228	Ab194228 Arabidops	c 905	33	70.2	668	13	ADU56526	Adu56526 Aspergill
c 833	33	70.2	452	13	AEA01269	Aea01269 Oryza spe	c 906	33	70.2	668	14	ADZ94529	Adz94529 Aspergill
c 834	33	70.2	452	13	AEA01284	Aea01284 Oryza spe	c 907	33	70.2	675	6	ABQ91724	Abq91724 M. capsul
c 835	33	70.2	452	13	AEA01280	Aea01280 Oryza spe	c 908	33	70.2	684	4	AAH08003	Aah08003 Human cDN
c 836	33	70.2	452	13	AEA01274	Aea01274 Oryza spe	c 909	33	70.2	691	3	AAFI2982	Aaf12982 Aspergill
c 837	33	70.2	452	13	AEA01283	Aea01283 Oryza spe	c 910	33	70.2	691	13	ADU57023	Adu57023 Aspergill
c 838	33	70.2	453	4	AAL23408	Aal23408 Human bre	c 911	33	70.2	691	14	ADZ95026	Adz95026 Aspergill
c 839	33	70.2	454	8	ABX93653	Abx93653 Human mic	c 912	33	70.2	712	12	ADO63102	Ado63102 Transcrip
c 840	33	70.2	457	13	AEA01317	Aea01317 Oryza spe	c 913	33	70.2	714	6	ABK75982	Abk75982 Bacillus
c 841	33	70.2	458	5	AAS83335	Aas83335 DNA enco	c 914	33	70.2	714	6	ABT09383	ABT09383 Phase-1 R
c 842	33	70.2	458	11	ACN90652	Acn90652 Breast ca	c 915	33	70.2	714	12	ADGA45435	Adg45435 Liver inf
c 843	33	70.2	458	13	AEA01319	Aea01319 Oryza spe	c 916	33	70.2	714	13	ADR31094	Adr31094 Spleen ne
c 844	33	70.2	460	4	AAL06064	Aal06064 Human rep	c 917	33	70.2	725	4	AAI95609	AAI95609 Human neu
c 845	33	70.2	460	4	ABL98629	Ab198629 Human tes	c 918	33	70.2	742	4	AAI97048	AAI97048 Human neu
c 846	33	70.2	460	13	AEA01316	Aea01316 Oryza spe	c 919	33	70.2	744	10	ADC91890	Adc91890 E. faeciu
c 847	33	70.2	463	9	ACH37677	Ach37677 Human end	c 920	33	70.2	746	14	ADZ48818	Adz48818 Insulin s
c 848	33	70.2	469	5	AAH83174	Aah83174 Human ova	c 921	33	70.2	750	4	AAI19425	AAI19425 Probe #93
c 849	33	70.2	473	4	ABAS8967	Abas8967 Human foe	c 922	33	70.2	750	4	ABA64438	ABA64438 Human foe
c 850	33	70.2	473	4	AAI38683	Aai38683 Probe #73	c 923	33	70.2	750	4	AAI44613	AAI44613 Probe #13
c 851	33	70.2	473	4	AAK32873	Aak32873 Human bon	c 924	33	70.2	750	4	ABA46575	ABA46575 Human bre
c 852	33	70.2	473	4	AAK07129	Aak07129 Human bra	c 925	33	70.2	750	4	ABA31574	ABA31574 Probe #10
c 853	33	70.2	473	4	ABS32599	Abs32599 Human liv	c 926	33	70.2	750	4	AAK38624	AAK38624 Human bon
c 854	33	70.2	473	6	ABS07677	Abs07677 Human gen	c 927	33	70.2	750	4	AAK12895	AAK12895 Human bra
c 855	33	70.2	473	9	ACH34737	Ach34737 Human end	c 928	33	70.2	750	5	AAH55843	Aah55843 Human liv
c 856	33	70.2	476	9	ACH27142	Ach27142 Human adu	c 929	33	70.2	750	5	AAH55843	Aah55843 Human SCN
c 857	33	70.2	477	3	AAPO7980	Aap07980 Fusarium	c 930	33	70.2	750	5	AAI05150	AAI05150 Probe #51
c 858	33	70.2	477	4	ABA26637	Abas26637 Probe #51	c 931	33	70.2	750	6	ABS12691	ABS12691 Human gen
c 859	33	70.2	477	4	AAK30645	Aak30645 Human bon	c 932	33	70.2	763	2	AAV65204	AAV65204 DNA enco
c 860	33	70.2	477	13	ADU52021	Adu52021 Fusarium	c 933	33	70.2	788	5	ADL45292	Adl45292 Human ova
c 861	33	70.2	477	14	ADZ90024	Adz90024 Fusarium	c 934	33	70.2	806	3	AAAC03773	AAc03773 Human sec
c 862	33	70.2	480	4	AAI83580	Aai83580 Human pol	c 935	33	70.2	813	10	ABX07889	Abx07889 S. pneumo
c 863	33	70.2	482	5	ABV45597	Abv45597 Human pro	c 936	33	70.2	813	12	ADM92013	Adm92013 S. pneumo
c 864	33	70.2	484	9	ACH41097	Ach41097 Human foe	c 937	33	70.2	816	13	ADK45027	Adk45027 Streptoco
c 865	33	70.2	486	12	ADQ18998	Adq18998 Human soi	c 938	33	70.2	819	4	AAH47035	Aah47035 Salmonell
c 866	33	70.2	490	9	ACH31329	Ach31329 Human adu	c 939	33	70.2	831	13	ADR92038	Adr92038 Novel S.
c 867	33	70.2	496	6	ABQ92838	Abq92838 Triticum	c 940	33	70.2	831	14	AEA55908	Aea55908 Streptoco
c 868	33	70.2	500	6	ABQ73178	Abq73178 C. glutam	c 941	33	70.2	834	10	ADH85451	Adh85451 Enterococ
c 869	33	70.2	500	6	ABK45135	Abk45135 cDNA enco	c 942	33	70.2	835	5	ABV29147	Abv29147 Human pro
c 870	33	70.2	502	2	AAV88792	AAV88792 EST clone	c 943	33	70.2	835	5	ABV23300	ABV23300 Human pro
c 871	33	70.2	504	10	ACD96659	AcD96659 Human col	c 944	33	70.2	844	4	AAAL20982	AAAL20982 Human bre
c 872	33	70.2	508	9	ACH29439	Ach29439 Human adu	c 945	33	70.2	852	4	AAK91599	AAK91599 Human cDN
c 873	33	70.2	515	13	ACF82315	Acf82315 Human SIR	c 946	33	70.2	852	12	ADL28026	AdL28026 5' end of
c 874	33	70.2	516	5	ADL62175	Adl62175 Human ova	c 947	33	70.2	852	12	ADL29598	AdL29598 5' end of
c 875	33	70.2	532	8	ABZ56067	Abz56067 Aspergill	c 948	33	70.2	864	10	ACF68524	ACF68524 Phototrab
c 876	33	70.2	541	13	ACN49002	Acn49002 Cotton pr	c 949	33	70.2	873	10	ABX06122	ABX06122 S. pneumo
c 877	33	70.2	544	5	ABV15688	Abv15688 Human pro	c 950	33	70.2	876	4	ABA08731	ABA08731 Prokaryot
c 878	33	70.2	550	14	ADX38737	Adx38737 Human can	c 951	33	70.2	876	8	ACA9771	ACA9771 Prokaryot
c 879	33	70.2	554	4	ABA08731	Abas08731 Human sec	c 952	33	70.2	876	11	ACN82310	ACN82310 Breast ca
c 880	33	70.2	560	12	ACH76124	Ach76124 Human gen	c 953	33	70.2				

954	33	70.2	876	13	ADK45509	Adk45509 Streptoco
955	33	70.2	883	6	ABK73423	Abk73423 Bacillus
956	33	70.2	883	13	ADK52485	Adk52485 Plant ful
957	33	70.2	894	13	ADK54092	Adk54092 Plant ful
958	33	70.2	894	14	ADK38778	Adk38778 Human can
959	33	70.2	897	13	ADR91945	Adr91945 Novel S.
960	33	70.2	897	14	AEA55815	Aea55815 Streptoco
961	33	70.2	906	4	AAF74632	Aaf74632 Bacillus
962	33	70.2	906	8	ADK63320	Adk63320 Human cDN
963	33	70.2	909	4	AAH53388	Aah53388 S. epider
964	33	70.2	917	13	ADR85532	Adr85532 Aspergill
965	33	70.2	918	6	ABQ90320	Abq90320 M. capaul
966	33	70.2	918	13	ADR62836	Adr62836 Cotton CD
967	33	70.2	939	6	ABN91620	Abn91620 Staphyloc
968	33	70.2	939	13	ADS03519	Ads03519 Staphyloc
969	33	70.2	940	6	ABK52047	Abk52047 cDNA enco
970	33	70.2	941	13	ADK27832	Adk27832 Plant ful
971	33	70.2	964	4	AAS26121	Aas26121 Human cDN
972	33	70.2	964	8	ABX73462	Abx73462 Human nov
973	33	70.2	975	13	ADT44085	Adt44085 Bacterial
974	33	70.2	977	13	ADR84945	Adr84945 Aspergill
975	33	70.2	983	13	ADR29192	Adr29192 Plant ful
976	33	70.2	984	4	AAK78603	Aak78603 Human imm
977	33	70.2	985	4	AAK78604	Aak78604 Human imm
978	33	70.2	985	4	AAK78602	Aak78602 Human imm
979	33	70.2	998	3	AACT77903	Aac77903 Human can
980	33	70.2	999	4	ABL06261	Ab106261 Drosophil
981	33	70.2	1005	13	ADS56465	Ads56465 Bacterial
982	33	70.2	1014	10	ADRF01403	Adf01403 Bacterial
983	33	70.2	1017	13	ADU25461	Adu25461 cDNA enco
984	33	70.2	1020	9	ADA29034	Ada29034 DNA enco
985	33	70.2	1031	13	ADQ86040	Adq86040 Human tum
986	33	70.2	1031	13	ADQ87091	Adq87091 Human tum
987	33	70.2	1076	4	AAS26542	Aas26542 Human cDN
988	33	70.2	1076	8	ABX73883	Abx73883 Human nov
989	33	70.2	1112	2	AAQL1412	Aaql1412 Rat cilia
990	33	70.2	1112	8	ABS57556	Abs57556 Human SEC
991	33	70.2	1129	14	ADK38776	Adk38776 Human can
992	33	70.2	1134	12	ADK16278	Adk16278 Nanoarcha
993	33	70.2	1137	14	ACL67294	Acl67294 M. xanthu
994	33	70.2	1140	4	ABL19455	Ab119455 Drosophil
995	33	70.2	1140	10	ADC92799	Adc92799 E. faeciu
996	33	70.2	1154	4	ABA07320	Abao7320 Human pan
997	33	70.2	1154	4	AAK90479	Aak90479 Human dig
998	33	70.2	1187	14	ADK38774	Adk38774 Human can
999	33	70.2	1209	13	ADR62837	Adr62837 Cotton CD
1000	33	70.2	1218	3	AAA37023	Aaa37023 Human PRO

ALIGNMENTS

RESULT 1	
ABT07721	ABT07721 standard; DNA; 927 BP.
ID	
XX	
AC	ABT07721;
XX	
DT	14-NOV-2002 (first entry)
XX	
DE	Breast cancer-associated gene sequence 29.
XX	
KW	Gene; ds; breast cancer; breast cancer-associated gene sequence;
KW	drug development; pharmacogenetics; biosensor development.
XX	
OS	Unidentified.
XX	
PN	WO200259377-A2.
XX	
PD	01-AUG-2002.
XX	
PF	24-JAN-2002; 2002WO-US002242.
XX	
PR	24-JAN-2001; 2001US-0263965P.
PR	
PR	02-FEB-2001; 2001US-0265928P.
PR	09-APR-2001; 2001US-00829472.
PR	09-APR-2001; 2001US-0282698P.
PR	04-MAY-2001; 2001US-0288590P.
XX	29-MAY-2001; 2001US-0294443P.
PA	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	
PI	Mack DH, Gish KC, Afar D;
XX	WPI; 2002-583738/62.
DR	N-PSDB; ABJ05564.
XX	
PT	Detecting a breast cancer-associated transcript in a patient's cell,
PT	useful for diagnosing breast cancer, comprises contacting a biological
PT	sample with a polynucleotide that selectively hybridizes with breast
XX	cancer nucleic acids.
PS	Claim 9; Page 372; 414pp; English.
XX	
CC	The invention comprises a method of detecting a breast cancer-associated
CC	transcript in a cell from a patient. The method of the invention involves
CC	contacting a biological sample from the patient with a nucleotide that
CC	hybridizes to one of the 69 breast cancer-associated gene sequences shown
CC	in the specification. The method of the invention is useful in the
CC	diagnosis or prognosis of breast cancer, and for detecting genes that are
CC	up or down-regulated in breast cancer cells. Genes identified by the
CC	method of the invention can be used in diagnostic purposes and also as
CC	targets for screening for therapeutic compounds that modulate breast
CC	cancer (e.g. hormones or antibodies). Identification of genes that are
CC	over or under expressed in breast cancer can additionally provide high-
CC	resolution, high-sensitivity datasets which can be used in the areas of
CC	diagnostics, therapeutics, drug development, pharmacogenetics, protein
CC	structure and biosensor development. DNA sequences ABT07693 - ABT07761
CC	represent the 69 breast cancer-associated gene sequences of the invention
SQ	Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	5.36 Length: 927
Score:	47.00 Matches: 9
Percent Similarity:	100.0% Conservative: 0
Best Local Similarity:	100.0% Mismatches: 0
Query Match:	100.0% Indels: 0
DB:	6 Gaps: 0
US-10-774-176-13 (1-9) x ABT07721 (1-927)	
Qy	1 PheLeuTyrLeuProArgAspValLeu 9
Db	
322 TTCCTTTACCTGCCGGGATGTGCTG 348	
RESULT 2	
ABX76333	
ID	ABX76333 standard; DNA; 927 BP.
XX	
AC	ABX76333;
XX	
DT	02-APR-2003 (first entry)
XX	
DE	Lung cancer-associated polynucleotide #197.
XX	
KW	Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW	antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW	chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX	interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX	
OS	Unidentified.
XX	
PN	WO200286443-A2.
XX	
PD	31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.
 XX PR 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (BOSB-) EOS BIOTECHNOLOGY INC.
 PA Aziz N, Murray R;
 XX WPI; 2003-093161/08.
 DR P-PSDB; ABUS6604.
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX Claim 22; Page 336; 453pp; English.
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.36 Length: 927
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-13 (1-9) x ABX76333 (1-927)
 QY 1 PheLeuTyrLeuProArgAspValLeu 9
 Db 322 TTCTTTTACCTGCGCGGGATGTGCTG 348
 RESULT 3
 ADB80503
 ID ADB80503 standard; DNA; 927 BP.
 XX ADB80503;
 AC ADB80503;
 DT 04-DEC-2003 (first entry)
 XX Ovarian cancer-associated transcript #34.
 DE cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 XX post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; de; gene.

XX OS Homo sapiens.
 XX Key Location/Qualifiers
 CDS 1..927
 /*tag= a
 WO2002102235-A2.
 XX 27-DEC-2002.
 XX 18-JUN-2002; 2002WO-US019297.
 XX 18-JUN-2001; 2001US-0299234P.
 PR 27-AUG-2001; 2001US-0315287P.
 PR 05-SEP-2001; 2001US-0317544P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (BOSB-) EOS BIOTECHNOLOGY INC.
 PA Mack DH, Gish KC;
 XX WPI; 2003-167431/16.
 DR P-PSDB; ADB80504.
 XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.
 XX Claim 10; Page 297; 332pp; English.
 XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancer, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selection of pre-cancerous lesions,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.
 XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.36 Length: 927
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-13 (1-9) x ADB80503 (1-927)
 QY 1 PheLeuTyrLeuProArgAspValLeu 9
 Db 322 TTCTTTTACCTGCGCGGGATGTGCTG 348
 RESULT 4
 ADB80503
 ID ADB80503 standard; cDNA; 927 BP.
 XX ADB80503;
 AC ADB80503;
 DT 17-JUN-2004 (first entry)
 XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
 DE
 XX

KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
XX vulnary; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
XX WO2003042661-A2.
XX 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
XX 21-NOV-2001; 2001US-0332464P.
XX 29-NOV-2001; 2001US-0334339P.
XX 03-DEC-2001; 2001US-0335394P.
XX 14-DEC-2001; 2001US-0340376P.
XX 08-JAN-2002; 2002US-0347211P.
XX 10-JAN-2002; 2002US-0347349P.
XX 08-FEB-2002; 2002US-0355250P.
XX 13-FEB-2002; 2002US-0356714P.
XX 20-FEB-2002; 2002US-0359077P.
XX 29-MAR-2002; 2002US-0368809P.
XX 04-APR-2002; 2002US-0370110P.
XX 12-APR-2002; 2002US-0372246P.
XX 05-JUN-2002; 2002US-0386614P.
XX 16-JUL-2002; 2002US-0396839P.
XX 22-JUL-2002; 2002US-0397775P.
XX 22-JUL-2002; 2002US-0397845P.
XX 09-SEP-2002; 2002US-0409450P.
XX (BOSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
XX P-PSDB; ADN38724.
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX Claim 8; SEQ ID NO 41; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or
XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX of determining the presence or absence of a pathological cell in a
XX patient by detecting a nucleic acid at least 80% identical to those of
XX the invention or by detecting a polypeptide of the invention. The
XX invention also relates to expression vectors and host cells comprising a
XX nucleic acid of the invention; antibodies which specifically bind a
XX polypeptide of the invention; use of such antibodies for drug targeting;
XX and methods of screening for modulators of activity or expression of the
XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX antibodies and methods are useful for diagnosing, prognosing and treating
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX neovascularisation syndromes, scarring and uterine fibroids. They may
XX also be useful in wound healing and in contraception. The present
XX sequence represents a nucleic acid sequence of the invention.
XX Sequence 927 BP; 187 A; 257 C; 240 G; 203 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.36 Length: 927
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0
US-10-774-176-13 (1-9) x ADN38723 (1-927)
Qy 1 PheLeuTyrlleuProArgAspValleu 9
Db 322 TTCTTTACCTGCGCGGATGCTG 348
RESULT 5
AAD56198
ID AAD56198 standard; DNA; 973 BP.
XX AAD56198;
XX 07-AUG-2003 (first entry)
XX Human LRRCAPS related DNA #5.
XX Human; p53 pathway; Leucine rich repeat capricious related protein;
XX LRRCAPS; cancer; gene therapy; ds.
XX Homo sapiens.
XX WO2003035831-A2.
XX 01-MAY-2003.
XX 21-OCT-2002; 2002WO-US033540.
XX 22-OCT-2001; 2001US-0338733P.
XX 15-FEB-2002; 2002US-0357600P.
XX 01-MAR-2002; 2002US-0361196P.
XX (EXEL-) EXELIXIS INC.
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
XX Francis-Lang H, Friedman L;
XX WPI; 2003-421410/39.
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
XX comprises contacting an assay system comprising a purified leucine rich
XX repeat, capricious related polypeptide or nucleic acid with a test agent.
XX Example 5; Page 74-75; 99pp; English.
XX The invention relates to a method of identifying a candidate p53 pathway
XX modulating agent. The method involves contacting an assay system
XX comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX polypeptide or nucleic acid or its fragment with a test agent and
XX detecting a test agent-biased activity, where a difference between the
XX test agent-biased activity and the reference activity identifies the test
XX agent as a candidate p53 pathway modulating agent. The method is useful
XX for identifying a candidate p53 pathway-modulating agent for preparing a
XX composition for diagnosing or treating cancer. The invention is useful in
XX gene therapy. The present sequence is human LRRCAPS related DNA
XX Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.67 Length: 973
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-10-774-176-13 (1-9) x AAD56198 (1-973)
Qy 1 PheLeuTyrlleuProArgAspValleu 9

Db 337 TTCTTTTACCTGCCGCGGATGCTG 363

RESULT 6

ABV99349

ID ABV99349 standard; DNA; 1156 BP.

AC ABV99349;

XX

DT 27-JAN-2003 (first entry)

XX

DE Human NOV8a coding sequence.

XX

KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;

KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;

KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;

KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;

KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;

KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;

KW immune disorder; haematopoietic disorder; cardiovascular disorder;

KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;

KW metabolic syndrome X; wasting disorder; cell differentiation; gene;

KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

XX

OS Homo sapiens.

XX

PN WO200272771-A2.

XX

PD 19-SEP-2002.

XX

PF 08-MAR-2002; 2002WO-US007288.

XX

PR 08-MAR-2001; 2001US-0274101P.

PR 08-MAR-2001; 2001US-0274194P.

PR 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274322P.

PR 08-MAR-2001; 2001US-0274849P.

PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.

PR 13-MAR-2001; 2001US-0275579P.

PR 13-MAR-2001; 2001US-0275601P.

PR 14-MAR-2001; 2001US-0276000P.

PR 16-MAR-2001; 2001US-0276776P.

PR 19-MAR-2001; 2001US-0276994P.

PR 20-MAR-2001; 2001US-0277239P.

PR 20-MAR-2001; 2001US-0277321P.

PR 20-MAR-2001; 2001US-0277327P.

PR 20-MAR-2001; 2001US-0277338P.

PR 21-MAR-2001; 2001US-0277791P.

PR 22-MAR-2001; 2001US-0277833P.

PR 23-MAR-2001; 2001US-0278152P.

PR 26-MAR-2001; 2001US-0278894P.

PR 27-MAR-2001; 2001US-0278999P.

PR 27-MAR-2001; 2001US-0279036P.

PR 28-MAR-2001; 2001US-0279344P.

PR 30-MAR-2001; 2001US-0279995P.

PR 30-MAR-2001; 2001US-0280233P.

PR 02-APR-2001; 2001US-0280802P.

PR 02-APR-2001; 2001US-0280822P.

PR 02-APR-2001; 2001US-0280900P.

PR 04-APR-2001; 2001US-0281194P.

PR 13-APR-2001; 2001US-0283675P.

PR 30-APR-2001; 2001US-0287424P.

PR 02-MAY-2001; 2001US-0288066P.

PR 03-MAY-2001; 2001US-0288342P.

PR 15-MAY-2001; 2001US-0288528P.

PR 16-MAY-2001; 2001US-0291190P.

PR 16-MAY-2001; 2001US-0291099P.

PR 30-MAY-2001; 2001US-0291240P.

PR 30-MAY-2001; 2001US-0294485P.

PR 31-MAY-2001; 2001US-0294889P.

PR 31-MAY-2001; 2001US-0294899P.

PR 18-JUN-2001; 2001US-0299037P.

PR 19-JUN-2001; 2001US-0299303P.

PR 19-JUN-2001; 2001US-0299310P.

PR 10-JUL-2001; 2001US-0304354P.

PR 31-JUL-2001; 2001US-0309198P.

PR 16-AUG-2001; 2001US-0312903P.

PR 10-SEP-2001; 2001US-0318462P.

PR 12-SEP-2001; 2001US-0318770P.

PR 27-SEP-2001; 2001US-0325430P.

PR 27-SEP-2001; 2001US-0325681P.

PR 18-OCT-2001; 2001US-0330380P.

PR 31-OCT-2001; 2001US-0335301P.

PR 14-NOV-2001; 2001US-0332172P.

PR 14-NOV-2001; 2001US-0332271P.

PR 14-NOV-2001; 2001US-0332272P.

PR 14-NOV-2001; 2001US-0333184P.

PR 14-NOV-2001; 2001US-0333272P.

PR 21-NOV-2001; 2001US-0332094P.

PR 03-DEC-2001; 2001US-0337426P.

PR 03-DEC-2001; 2001US-0338092P.

PR 04-DEC-2001; 2001US-0337185P.

PR 03-JAN-2002; 2002US-0345705P.

PR 08-MAR-2002; 2002US-00093463.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;

PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CNM;

PI Pena CRA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;

PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;

PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;

PI Zhong N;

XX

DR WPI; 2002-732824/79.

DR P-PSDB; ABP70071.

XX

PT New NOVX polypeptides and polynucleotides, useful for preventing,

PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,

PT Alzheimer's disease, dyslipidemia, obesity, immune or hematopoietic

PT disorders, and asthma.

XX

PS Claim 16; Page 114-115; 619pp; English.

XX

CC The present invention relates to new isolated proteins (NOVX) and their

CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is

CC any number from 1 to 48. The NOVX proteins and coding sequences are

CC useful in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, preferably a NOVX-associated disorder.

CC The NOVX coding sequences and proteins are useful for treating,

CC preventing or diagnosing diseases such as metabolic disorders, diabetes,

CC obesity, infectious disease, anorexia, cancer-associated cachexia,

CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's

CC disease, immune disorders, haematopoietic disorders, cardiovascular

CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic

CC disturbances associated with obesity, metabolic syndrome X or wasting

CC disorders associated with chronic diseases or various cancers. The NOVX

CC coding sequences and proteins may also be used as targets for the

CC identification of small molecules that modulate or inhibit e.g.

CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,

CC wound healing and angiogenesis, in gene therapy, in generation of

CC antibodies that bind immunospecifically to NOVX substances for use in

CC therapeutic or diagnostic methods

XX

SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.93	Length:	1156
Score:	47.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-13 (1-9) x ABV99349 (1-1156)

```
QY      1 PheLeuTyrLeuProArgAspValLeu 9
Db      553 TTCTTTACCTGCCGCGGATGCTG 579

RESULT 7
AA27058
ID AAA27058 standard; DNA; 1263 BP.
XX
AC AAA27058;
XX
XX 22-AUG-2000 (first entry)
DT
DE Human 5T4 tumour-associated antigen gene.
XX
XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
XX Homo sapiens.
OS
XX WO200029428-A2.
PN
XX 25-MAY-2000.
PD
XX 18-NOV-1999; 99WO-GB003859.
PF
XX 18-NOV-1998; 98GB-00025303.
PR
XX 27-JAN-1999; 99GB-00001739.
PR
XX 30-JUL-1999; 99GB-00017995.
PR
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX Carroll MW, Myers KA;
XX WPI; 2000-387735/33.
XX
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
PT response useful in vaccinating against and in treating tumors.
XX
XX Example 2; Page 78; 79pp; English.
XX
XX The present sequence encodes the human 5T4 tumour-associated antigen
CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
CC carcinomas but has a highly restricted expression pattern in normal adult
CC tissues. It appears to be strongly correlated to metastasis in colorectal
CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
CC induced were inoculated with a virus expression vector containing the
CC present sequence. The 5T4 antigen was shown to be effective at eliciting
CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
CC the antigen and the antigen itself can be used to elicit an immune
CC response, preferably CTL or an antibody response in a subject
XX
SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.67 Length: 1263
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-13 (1-9) x AAA27058 (1-1263)

QY      1 PheLeuTyrLeuProArgAspValLeu 9
Db      664 TTCTTTACCTGCCGCGGATGCTG 690

RESULT 8
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX
```

```
AC      AAF89736;
XX
DT      23-JUL-2001 (first entry)
XX
DE      Nucleotide sequence of canine 5T4 protein.
XX
KW      Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
KW hypersensitivity; autoimmune disease; central nervous system disorder;
KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
KW cardiovascular disease; gastrointestinal disease; infection; diabetes;
KW Helicobacter-related disease; immune disorder; ss.
XX
OS      Canis sp.
XX
XX Key Location/Qualifiers
FH      1..1263
CDS     /*tag= a
FT      /product= "5T4"
FT
FT
XX      WO200136486-A2.
XX
XX 25-MAY-2001.
PD
XX 13-NOV-2000; 2000WO-GB004317.
PF
XX 18-NOV-1999; 99WO-GB003859.
PR
XX 15-FEB-2000; 2000GB-00003527.
PR
XX 02-MAR-2000; 2000GB-00005071.
PR
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
XX Myers KA;
XX WPI; 2001-343805/36.
XX P-PSDB; AAB83839.
XX
XX Use of single chain antibody capable of recognizing a disease associated
PT molecule for manufacturing a medicament for preventing and/or treating a
PT disease condition associated with disease associated molecule.
XX
XX Disclosure; Fig 26; 118pp; English.
XX
XX The specification describes the use of a single chain antibody (ScFv),
CC which is capable of recognizing a disease associated molecule in the
CC manufacture of a medicament for the prevention and treatment of a disease
CC condition. The ScFv antibody is useful in the manufacture of a
CC medicament, for affecting a disease in vivo, for preparing a
CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
CC treatment of a disease. The ScFv antibody is also useful for treating
CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
CC diseases, cancers, central nervous system disorders including Parkinson's
CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
CC related diseases, and other immune disorders. The present sequence
CC encodes a 5T4 protein, which is used to produce ScFv of the invention
XX
SQ      Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.67 Length: 1263
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-13 (1-9) x AAF89736 (1-1263)

QY      1 PheLeuTyrLeuProArgAspValLeu 9
Db      664 TTCTTTACCTGCCGCGGATGCTG 690
```

RESULT 9
 ABK87174
 ID ABK87174 standard; cDNA; 1263 BP.
 XX AC
 XX ABK87174;
 DT 07-OCT-2002 (first entry)
 XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 574.
 DE Canine; dog; oncofoetal leucine-rich glycoprotein; 574; tumour;
 XX cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX Canis sp.
 OS
 XX Key Location/Qualifiers
 PH CDS 1..1263
 FT /*tag= a
 FT /product= "574 protein"
 XX WO200238612-A2.
 PN 16-MAY-2002.
 XX
 XX 13-NOV-2001; 2001WO-GB005004.
 XX 13-NOV-2000; 2000WO-GB004317.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Myers K, Drury N, Carroll M;
 XX WPI; 2002-557449/59.
 DR P-PSDB; AAU98693.
 XX Novel canine or feline 574 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.
 XX Claim 1; Page 67; 68pp; English.
 PS The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 574, and the
 CC polynucleotide sequences encoding them. The 574 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 574 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 574
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus (es).
 CC The present sequence encodes canine 574 protein
 XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7.67 Length: 1263
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-13 (1-9) x ABK87174 (1-1263)

QY 1 PheLeuTyrLeuProArgAspValLeu 9
 Db 664 TTCCCTTACTTGCCTCGCGACGTCCTG 690
 RESULT 10
 AAD56199
 ID AAD56199 standard; DNA; 1331 BP.
 XX AC
 XX AAD56199;
 DT 07-AUG-2003 (first entry)
 XX Human LRRCAPS related DNA #6.
 DE Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX Homo sapiens.
 OS
 XX WO2003035831-A2.
 FN 01-MAY-2003.
 PD 21-OCT-2002; 2002WO-US033540.
 XX 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX (EXEL-) EXELIXIS INC.
 XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 PI WPI; 2003-421410/39.
 DR Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX Disclosure; Page 75-76; 99pp; English.
 PS The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8.15 Length: 1331
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-13 (1-9) x AAD56199 (1-1331)

QY 1 PheLeuTyrLeuProArgAspValLeu 9
 Db 694 TTCCCTTACTTGCCTCGCGACGTCCTG 720
 RESULT 11
 ADJ56299
 ID ADJ56299 standard; cDNA; 2020 BP.
 XX

AC ADJ56299;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.
XX
KW human, differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX
OS Homo sapiens.
XX
PN US2003119009-A1.
XX
PD 26-JUN-2003.
XX
PF 25-FEB-2002; 2002US-00084817.
XX
PR 23-FEB-2001; 2001US-0270784P.
XX
PA (STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLOW/) PLOW S E.
PA (SHOH/) SHOHET J M.
XX
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX
DR WPI; 2003-635698/60.
XX
XX New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
PS Claim 1; SEQ ID NO 105; 27pp; English.
XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?docID=20030119009.
XX
SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 13.2 Length: 2020
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-13 (1-9) x ADJ56299 (1-2020)
Qy 1 PheLeuTyrLeuProArgAspValLeu 9
Db 734 TTCCTTTACTGCGCGGGATGCTG 760
RESULT 12
ACCS1052
ID ACCS1052 standard; cDNA; 2053 BP.

XX ACCS1052;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated cDNA sequence SEQ ID NO:192.
XX
KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003003906-A2.
XX
PD 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US021338.
XX
PR 03-JUL-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Aziz N;
XX
DR WPI; 2003-201532/19.
DR P-PSDB; ABR48236.
XX
XX Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
PS Claim 6; Page 296; 307pp; English.
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridises to a sequence that is 80 % identical to a
CC table of sequences (see ACCS0951 to ACCS1059). ACCS0951 to ACCS1059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 13.5 Length: 2053
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-10-774-176-13 (1-9) x ACCS1052 (1-2053)
Qy 1 PheLeuTyrLeuProArgAspValLeu 9
Db 748 TTCCTTTACTGCGCGGGATGCTG 774
RESULT 13
ABX76332
ID ABX76332 standard; DNA; 2053 BP.
XX
AC ABX76332;

XX 02-APR-2003 (first entry)
 XX Lung cancer-associated polynucleotide #196.
 XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 XX WO200286443-A2.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012476.
 XX 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (BOSB-) EOS BIOTECHNOLOGY INC.
 XX Aziz N, Murray R;
 XX WPI; 2003-093161/08.
 DR P-PSDB; ABUS6603.
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX Claim 22; Page 335; 453pp; English.
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores: Pred. No.: 13.5 Length: 2053
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-13 (1-9) x ABX76332 (1-2053)
 QY 1 PheLeuTyRLeuProArgAspValLeu 9
 DB 748 TTCTTTTACCTGCGCGGATGTGCTG 774
 RESULT 15
 AAD56200
 ID AAD56200 standard; DNA; 2053 BP.
 XX
 AC AAD56200;

Db 748 TTCTTTTACCTGCGCGGATGTGCTG 774
 RESULT 14
 AAD56197
 ID AAD56197 standard; DNA; 2053 BP.
 XX
 AC AAD56197;
 XX 07-AUG-2003 (first entry)
 XX Human LRRCAPS DNA #11.
 XX Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX Homo sapiens.
 XX WO2003035831-A2.
 XX 01-MAY-2003.
 XX 21-OCT-2002; 2002WO-US033540.
 PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX (EXEL-) EXELIXIS INC.
 XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX WPI; 2003-421410/39.
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX Example 5; Page 73-74; 99pp; English.
 XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores: Pred. No.: 13.5 Length: 2053
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-13 (1-9) x AAD56197 (1-2053)
 QY 1 PheLeuTyRLeuProArgAspValLeu 9
 DB 748 TTCTTTTACCTGCGCGGATGTGCTG 774
 RESULT 15
 AAD56200
 ID AAD56200 standard; DNA; 2053 BP.
 XX
 AC AAD56200;

XX 07-AUG-2003 (first entry)
 XX Human LRRCAPS DNA #12.
 XX Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX Homo sapiens.
 OS
 XX WO2003035831-A2.
 PN
 XX 01-MAY-2003.
 PD
 XX 21-OCT-2002; 2002WO-US033540.
 XX
 PF
 XX 22-OCT-2001; 2001US-0338733P.
 PR
 XX 15-FEB-2002; 2002US-0357600P.
 PR
 XX 01-MAR-2002; 2002US-0361196P.
 PR
 XX (EXEL-) EXELIXIS INC.
 PA
 XX Belvin M. Schlieithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 PI WPI; 2003-421410/39.
 XX
 DR Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 XX Disclosure; Page 76-77; 99pp; English.
 XX
 XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC testing a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 13.5 Length: 2053
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-13 (1-9) x AAD56200 (1-2053)
 QY 1 PheLeuTyrLeuProArgAspValLeu 9
 Db 748 TTCTTTACCTGCGCGGARGTCTG 774
 RESULT 16
 ID ADN38721 standard; cDNA; 2053 BP.
 XX
 AC ADN38721;
 XX
 XX 17-JUN-2004 (first entry)
 DT
 XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.
 DE
 XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW

retinal neovascularisation syndrome; scarring; uterine fibroid;
 detection; diagnosis; prognosis; drug screening; drug targeting;
 wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 vulnery; gene therapy; vaccine; gene; ss.
 Homo sapiens.
 OS
 XX WO2003042661-A2.
 PN
 XX 22-MAY-2003.
 PD
 XX 13-NOV-2002; 2002WO-US036810.
 XX
 PF
 XX 13-NOV-2001; 2001US-0350666P.
 PR
 XX 21-NOV-2001; 2001US-0332464P.
 PR
 XX 29-NOV-2001; 2001US-0334393P.
 PR
 XX 03-DEC-2001; 2001US-0335394P.
 PR
 XX 14-DEC-2001; 2001US-0340376P.
 PR
 XX 08-JAN-2002; 2002US-0347211P.
 PR
 XX 10-JAN-2002; 2002US-0347349P.
 PR
 XX 08-FEB-2002; 2002US-0355250P.
 PR
 XX 13-FEB-2002; 2002US-0356714P.
 PR
 XX 20-FEB-2002; 2002US-0359077P.
 PR
 XX 29-MAR-2002; 2002US-0358809P.
 PR
 XX 04-APR-2002; 2002US-0370110P.
 PR
 XX 12-APR-2002; 2002US-0372246P.
 PR
 XX 05-JUN-2002; 2002US-0386614P.
 PR
 XX 16-JUL-2002; 2002US-0396839P.
 PR
 XX 22-JUL-2002; 2002US-0397775P.
 PR
 XX 22-JUL-2002; 2002US-0397845P.
 PR
 XX 09-SEP-2002; 2002US-0409450P.
 PR
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 PI WPI; 2003-468649/44.
 DR P-PSDB; ADN38722.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 XX Claim 8; SEQ ID NO 39; 1385pp; English.
 XX
 XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 13.5 Length: 2053
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

US-10-774-176-13 (1-9) x ADN38721 (1-2053)

QY 1 PheLeuTyrLeuProArgAspValIeu 9

Db 748 TTCCTTTACCTGCCGCGGATGTGCTG 774

RESULT 17

ADL06473
ID ADL06473 standard; cDNA; 2053 BP.

AC ADL06473;

XX 20-MAY-2004 (first entry)

DT Human tumour-associated antigenic target (TAT) cDNA sequence #53.

DE Human; tumour-associated antigenic target; TAT; cell death; tumour;

KW cancer; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004016225-A2.

XX 26-FEB-2004.

XX 19-AUG-2003; 2003WO-US025892.

XX 19-AUG-2002; 2002US-0404809P.

PR 21-AUG-2002; 2002US-0405645P.

PR 23-SEP-2002; 2002US-0413192P.

PR 15-OCT-2002; 2002US-0419008P.

PR 15-NOV-2002; 2002US-0426847P.

PR 02-JUL-2003; 2003US-0484959P.

XX (GETH) GENENTECH INC.

PI Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Poleon A, Smith V;

PI Spencer SD, Wu TD, Zhang Z;

XX WPI; 2004-257144/24.

DR P-PSDB; ADL06552.

PT New antibody that binds to a tumor-associated antigenic target (TAT)
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating cancer.

XX Claim 1; SEQ ID NO 53; 319pp; English.

XX The present invention relates to the isolation of human tumour-associated
CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also
CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is
CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a
CC humanised antibody. It is conjugated to a growth inhibitory agent. It is
CC produced in bacteria or in CHO cells and induces death of a cell to which
CC it binds. The antibody is useful for preparing a composition for
CC diagnosing or treating tumours and cancer. The present sequence
CC represents a human TAT cDNA sequence of the invention.

XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 13.5 Length: 2053
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-13 (1-9) x ADL06473 (1-2053)

QY 1 PheLeuTyrLeuProArgAspValIeu 9

Db 748 TTCCTTTACCTGCCGCGGATGTGCTG 774

Db 748 TTCCTTTACCTGCCGCGGATGTGCTG 774

RESULT 18

ADN03961

ID ADN03961 standard; cDNA; 2053 BP.

XX AC ADN03961;

XX 01-JUL-2004 (first entry)

DT Antipsoriatic cDNA sequence #180.

DE ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

XX WO2004028479-A2.

XX 08-APR-2004.

XX 25-SEP-2003; 2003WO-US030907.

XX 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.

PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;

XX WPI; 2004-305105/28.

DR P-PSDB; ADN03962.

PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.

XX Claim 1; SEQ ID NO 355; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 13.5 Length: 2053
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-13 (1-9) x ADN03961 (1-2053)

QY 1 PheLeuTyrLeuProArgAspValIeu 9

Db 748 TTCCTTTACCTGCCGCGGATGTGCTG 774

RESULT 19

ADR25444

ID ADR25444 standard; DNA; 2053 BP.

XX AC ADR25444;

XX 21-OCT-2004 (first entry)

DE Breast cancer prognosis marker #1305.

XX ds; breast cancer; prognosis; gene expression; diagnosis.

XX Homo sapiens.

XX PD 22-MAY-2003.
XX PF 18-JUN-2002; 2002US-00175523.
XX PR 18-JUN-2001; 2001US-0299151P.
XX PR 07-SEP-2001; 2001US-0317828P.
XX PR 25-SEP-2001; 2001US-0325150P.
XX PR 14-NOV-2001; 2001US-0333047P.
XX PR 18-JAN-2002; 2002US-0349936P.
XX PR 04-MAR-2002; 2002US-0361834P.
XX PA (PSYC-) PSYCHIATRIC GENOMICS INC.
XX PI Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
XX PI Palfreyman M, Rajan P;
XX PI WPI; 2004-118903/12.
XX DR Identifying a compound that can treat disease or disorders, such as, a
XX PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
XX PT determining the expression of one or more efficacy genes in a cell
XX PT contacted with the test compound.
XX PS Example 6; SEQ ID NO 174; 39pp; English.
XX CC This invention relates to a novel screening method identified as a multi-
XX CC parameter high throughput screening (MPHTS) assay. Specifically, it
XX CC refers to an assay that utilises the disease signature of a plurality of
XX CC specific genes associated with a particular disease, and identifies
XX CC differential expression between those cells taken from individuals
XX CC affected by that disease and those that are not affected. The present
XX CC invention then describes the screening of candidate pharmaceutical
XX CC compounds to identify those that have a potential therapeutic benefit for
XX CC the treatment of neuropsychiatric and neurodegenerative disorders
XX CC including schizophrenia, bipolar affective disorder (BAD) and autism, as
XX CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
XX CC of this invention exhibit various activities including neuroleptic,
XX CC nootropic, antianic and antidepressant. Furthermore, the screening
XX CC method used in MPHTS will be automated, such that a large number of test
XX CC compounds may be rapidly screened with a minimal amount of labour and
XX CC effort. This polynucleotide is a human cDNA sequence of a gene that is
XX CC differentially expressed in the presence of a therapeutic compound and
XX CC represents an exemplary efficacy gene for bipolar affective disorder,
XX CC given in an exemplification of the invention.
XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 13.5 Length: 2053
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 13 Indels: 0
DB: Gaps: 0
US-10-774-176-13 (1-9) x ADV35098 (1-2053)
Qy 1 PheLeuTyrLeuProArgAspValleu 9
Db 748 TTCCTTTACCTGCGCGGATGTGCTG 774
RESULT 22
AAS87175
ID AAS87175 standard; cDNA; 2338 BP.
XX AC AAS87175;
XX XX
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #22979.
XX XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX DR P-PSDB; ABG222388.
XX DR New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 22979; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 15.7 Length: 2338
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0
US-10-774-176-13 (1-9) x AAS87175 (1-2338)
Qy 1 PheLeuTyrLeuProArgAspValleu 9
Db 1005 TTCCTTTACCTGCGCGGATGTGCTG 1031
RESULT 23
AAK94253
ID AAK94253 standard; cDNA; 2359 BP.
XX AC AAK94253;
XX XX
XX DT 06-NOV-2001 (first entry)
XX XX

```

DE Human full-length cDNA, SEQ ID NO: 2864.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR P-PSDB; AAM93333.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5' and 3' ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
CC from EPO
XX
SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 15.8 Length: 2359
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-13 (1-9) x AAK94253 (1-2359)
Qy 1 PheLeuTyrlleuProArgAspValleu 9
Db 1087 TTCCTTTACCTGCGCGGATGCTG 1113

RESULT 24
ADL30831
ID ADL30831 standard; cDNA; 2359 BP.
XX
AC ADL30831;
XX
XX 20-MAY-2004 (first entry)
DT
DE Full length human cDNA clone SeqID 2864.
XX
KW human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss; gene.
XX
OS Homo sapiens.
XX

```

```

PN EP1396543-A2.
XX
PD 10-MAR-2004.
XX
PF 07-JUL-2000; 2003EP-00025638.
XX
PX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
DR P-PSDB; ADL30832.
XX
PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
PS Example 1; SEQ ID NO 2864; 1340pp; English.
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction.
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX
SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 15.8 Length: 2359
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-13 (1-9) x ADL30831 (1-2359)
Qy 1 PheLeuTyrlleuProArgAspValleu 9
Db 1087 TTCCTTTACCTGCGCGGATGCTG 1113

RESULT 25
AAK94254
ID AAK94254 standard; cDNA; 2361 BP.
XX
AC AAK94254;
XX
XX 06-NOV-2001 (first entry)
DT
DE Human full-length cDNA, SEQ ID NO: 2866.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX

```

XX (HELI-) HELIX RES INST.
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR P-PSDB; AAM93334.
 XX
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 PT
 XX Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.
 PS
 XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from BPO
 XX
 SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.8 Length: 2361
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-13 (1-9) x AAK94254 (1-2361)
 QY 1 PhleuTyLeuProArgAspValleu 9
 Db 1089 TTCTTTTACCTGCCGCGGATGTGCTG 1115
 RESULT 26
 ADI26162
 ID ADI26162 standard; cDNA; 2361 BP.
 XX
 AC ADI26162;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #64.
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 XX 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00137505.

PR 16-MAY-2003; 2003US-0470836P.
 XX
 PA (ASAH) ASahi KASEI KK.
 XX
 PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX
 DR WPI; 2004-122214/12.
 DR P-PSDB; ADI26163.
 XX
 XX New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 PT
 XX Claim 4; SEQ ID NO 127; 1368pp; English.
 PS
 XX The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 15.8 Length: 2361
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-13 (1-9) x ADI26162 (1-2361)
 QY 1 PhleuTyLeuProArgAspValleu 9
 Db 1089 TTCTTTTACCTGCCGCGGATGTGCTG 1115
 RESULT 27
 ADL30833
 ID ADL30833 standard; cDNA; 2361 BP.
 XX
 AC ADL30833;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Full length human cDNA clone SeqID 2866.
 XX
 KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.
 XX
 OS Homo sapiens.
 XX
 PN EP1396543-A2.
 XX
 XX 10-MAR-2004.

```

XX
PP 07-JUL-2000; 2003EP-00025638.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183865.
XX
PR 07-JUL-2000; 2000EP-00114089.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2004-204755/20.
DR P-PSDB; ADL30834.
XX
PT New oligonucleotide primers (830 CDNAS) useful for synthesizing full
PT length human CDNAS.
XX
PS Example 1; SEQ ID NO 2866; 1340pp; English.
XX
CC This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX
SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 15.8 Length: 2361
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-13 (1-9) x ADL30833 (1-2361)
QY 1 PheLeuTyrlLeuProArgAspValleu 9
Db 1089 TTCTTTACCTGCGCGGGATGCTG 1115

RESULT 28
ABK87175
ID ABK87175 standard; cDNA; 1260 BP.
XX
AC ABK87175;
XX
DT 07-OCT-2002 (first entry)
XX
DE cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
XX
KW Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX
OS Felis sp.
XX
FH Key Location/Qualifiers
FT CDS 1..1260
FT /tag= a
FT /product= "5T4 protein"
XX
PN WO200238612-A2.
XX
PD 16-MAY-2002.
XX

```

```

PF 13-NOV-2001; 2001WO-GB005004.
XX
PR 13-NOV-2000; 2000WO-GB004317.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Myers K, Drury N, Carroll M;
XX
XX WPI; 2002-557449/59.
DR P-PSDB; AAU98694.
XX
PT Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 4; Page 68; 68pp; English.
XX
CC The present invention relates to the isolation of canine and feline
CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 5T4 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 5T4
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus(es).
CC The present sequence encodes feline 5T4 protein
XX
SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 52.6 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-13 (1-9) x ABK87175 (1-1260)
QY 1 PheLeuTyrlLeuProArgAspValleu 9
Db 661 TTCTCTTCTTCTGCGGACGTACTG 687

RESULT 29
ADB97513
ID ADB97513 standard; DNA; 1260 BP.
XX
AC ADB97513;
XX
DT 04-DEC-2003 (first entry)
XX
DE Feline 5T4 antigen DNA.
XX
KW Major Histocompatibility Complex class I peptide epitope; MHC;
KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
KW cytostatic; cancer; feline; gene; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 1..1260
FT /tag= a
FT /product= "feline 5T4 antigen protein"
XX
PN WO2003068816-A1.

```



```

XX PD 21-AUG-2003.
XX PF 13-FEB-2003; 2003WO-GB000670.
XX PF 13-FEB-2002; 2002GB-00003419.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll M, Kingsman S, Redchenko I;
XX DR WPI; 2003-637141/60.
XX DR P-PSDB; ADB97520.
XX PT New major histocompatibility complex class I peptide epitopes from human
XX PT 574 tumor-associated antigen, useful for preventing and/or treating a
XX PT disease, particularly cancer.
XX PS Disclosure; Page 67; 73pp; English.
XX CC The invention relates to a novel Major Histocompatibility Complex (MHC)
XX CC class I peptide epitope of the 574 antigen. The invention further
XX CC provides a polypeptide string comprising the 574 epitope; a nucleic acid
XX CC sequence encoding the 574 epitope or a polypeptide string of the 574
XX CC epitope; a vector system capable of delivering the 574 epitope nucleic
XX CC acid to a cell; a cell pulsed with the 574 epitope, a polypeptide of the
XX CC 574 epitope, its encoding nucleic acid, or the vector system; a vaccine
XX CC comprising the above; a method for treating and/or preventing a disease
XX CC in a subject by administering the vaccine; an agent capable of binding
XX CC specifically to the 574 epitope and/its encoding nucleic acid; a method
XX CC comprising detecting the presence of the 574 epitope or its encoding
XX CC nucleic acid in a subject; and a T cell line or clone capable of
XX CC specifically recognising the 574 epitope in conjunction with an MHC class
XX CC I molecule. The 574 epitope has cytostatic activity. The vaccine
XX CC comprising the 574 epitope or its encoding nucleic acid and the vector
XX CC system or cell is useful in the prevention and/or treatment of a disease,
XX CC particularly cancer. The detection method is useful for diagnosing or
XX CC monitoring the progression of a cancerous disease, and for detecting the
XX CC presence of the 574 epitope or its nucleic acid. The T cell line or clone
XX CC is useful in the manufacture of a medicament for treating and/or
XX CC preventing a disease. This polynucleotide sequence represents the feline
XX CC 574 antigen coding DNA of the invention.
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 52.6 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-13 (1-9) x ADB97513 (1-1260)
QY 1 PheLeuTyLeuProArgAspValleu 9
Db 661 TTCTCTTCTTGCCTCGGACGTACTG 687

RESULT 30
ADB97452
ID ADB97452 standard; DNA; 1260 BP.
XX AC ADB97452;
XX DT 04-DEC-2003 (first entry)
XX DE DNA encoding feline 574 protein.
XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
XX KW epitope; 574 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
XX OS Unidentified.

XX Key Location/Qualifiers
XX CDS 1..1260
XX FT /*tag= a
XX FT /product= "Feline 574 antigen protein"
XX PN WO2003068815-A2.
XX XX 21-AUG-2003.
XX XX 13-FEB-2003; 2003WO-GB000618.
XX XX 13-FEB-2002; 2002GB-00003420.
XX XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX XX Carroll M, Harrop R, Kingsman S;
XX XX WPI; 2003-663795/62.
XX XX P-PSDB; ADB97455.
XX PT New Major Histocompatibility Complex class II peptide epitope of 574,
XX PT useful for manufacturing a medicament for diagnosing, preventing and/or
XX PT treating a disease, e.g. cancer.
XX PS Disclosure; Page 49; 63pp; English.
XX CC The invention relates to a Major Histocompatibility Complex (MHC) class
XX CC II peptide epitope of the 574 antigen. The vaccine or T-cell line or
XX CC clone has a cytostatic activity, as it is useful in manufacturing a
XX CC medicament for preventing and/or treating a disease, particularly cancer.
XX CC The methods are useful for detecting T-cells capable of specifically
XX CC recognising a peptide epitope in conjunction with an MHC molecule, for
XX CC diagnosing or monitoring the progression of a cancerous disease, or for
XX CC detecting the presence of a peptide or nucleic acid using an agent. The
XX CC MHC class II peptide epitope of the invention can be used in gene therapy
XX CC or as part of a vaccine. This polynucleotide sequence represents the DNA
XX CC coding for the feline 574 protein.
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 52.6 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-13 (1-9) x ADB97452 (1-1260)
QY 1 PheLeuTyLeuProArgAspValleu 9
Db 661 TTCTCTTCTTGCCTCGGACGTACTG 687

RESULT 31
ADC00087/c
ID ADC00087 standard; DNA; 86248 BP.
XX AC ADC00087;
XX DT 04-DEC-2003 (first entry)
XX DE Enterohaemorrhagic E. coli 0157:H7-specific nucleic acid SEQ ID NO: 132.
XX KW ds; gene; enterohaemorrhagic; anti-bacterial.
XX OS Escherichia coli; 0157:H7.
XX PN JP2002355074-A.
XX PD 10-DEC-2002.
XX XX

```

```

PP 24-JAN-2002; 2002JP-00015959.
XX
PR 24-JAN-2001; 2001JP-00112010.
XX
PA (UYTS-) UNIV TSUKUBA.
XX
XX WPI; 2003-451640/43.
XX
XX Enterohemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
XX Claim 2; SEQ ID NO 132; 2067pp; Japanese.
XX
XX The invention relates to a novel enterohaemorrhagic Escherichia coli
CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC has anti-bacterial activity. The polypeptide can be used in detection
CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present
CC sequence represents an E. coli O157:H7-specific nucleic acid of the
CC invention.
XX
SQ Sequence 86248 BP; 22338 A; 20003 C; 21303 G; 22604 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.04e+03 Length: 86248
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-13 (1-9) x ADC00087 (1-86248)

Qy 1 PheLeuTyrLeuProArgAspValLeu 9
Db 64602 TTCTTTATCGCTAAGGATATTC 64576

RESULT 32
ACD19044/c
ID ACD19044 standard; DNA; 87563 BP.
XX ACD19044;
XX
XX 27-OCT-2003 (revised)
DT 21-AUG-2003 (first entry)
XX
DE E. coli O157 unique DNA sequence OZID_57.
XX
XX OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
KW food poisoning.
XX
XX Escherichia coli; strain O157:H7.
XX
XX US2003023075-A1.
PN
XX 30-JAN-2003.
PD
XX 01-APR-2002; 2002US-00114170.
PF
XX 04-DEC-1998; 98US-0110955P.
PR 03-DEC-1999; 99US-00453702.
XX
XX (BLAT/) BLATTNER F R.
PA (BURL/) BURLAND V D.
PA (PERN/) PERNA N T.
PA (PLUN/) PLUNKETT G.
PA (WELC/) WELCH R.
XX
XX Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;
PI WPI; 2003-479497/45.
XX
XX New DNA sequences from Escherichia coli strain O157:H7, useful for
PT

```

```

PT detecting E. coli O157:H7 in a sample, or in designing diagnostic probes
PT which can be used to distinguish strain O157:H7 from strain K12 using
XX molecular techniques.
XX
XX Claim 16; SEQ ID NO 57; 33pp; English.
XX
XX The invention relates to an isolated DNA molecule comprising an E. coli
CC strain O157:H7 sequence selected from a clostridial cytotoxin-like gene,
CC a urease gene cluster, a KIX toxin-like gene cluster, a locus of
CC enterocyte effacement and 2 genes from its associated lymphocytic phage
CC 933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
CC O157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC (which can develop into haemolytic uraemic syndrome). Also included are
CC an isolated DNA molecule comprising a nucleotide sequence identical to at
CC least 25 contiguous nucleotides contained in DNA sequences selected from
CC ACD1898-ACD19242 (being 255 E.coli O157 DNA sequences which are not
CC found in E.coli K12), a recombinant DNA construction comprising the DNA
CC above and a method for detecting E. coli O157:H7 (ATCC 43895) in a sample
CC (or distinguishing between O157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC O157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with O157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain O157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli O157:H7 DNA
CC sequence (termed OZID_1-OZID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030023075 (Updated on 27-OCT-2003
CC to standardise OS field)
XX
SQ Sequence 87563 BP; 22620 A; 20384 C; 21612 G; 22935 T; 0 U; 12 Other;

Alignment Scores:
Pred. No.: 7.17e+03 Length: 87563
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-13 (1-9) x ACD19044 (1-87563)

Qy 1 PheLeuTyrLeuProArgAspValLeu 9
Db 64590 TTCTTTATCGCTAAGGATATTC 64564

RESULT 33
AAA27060
ID AAA27060 standard; DNA; 901 BP.
XX AAA27060;
XX
XX 22-AUG-2000 (first entry)
DT
XX
XX Canine 5T4 tumour-associated antigen gene.
XX
XX Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
XX Canis sp.
XX
XX Key Location/Qualifiers
FT CDS 1..858
FT /tag= a
FT /product= "5T4 antigen"
FT misc_feature 61..74
FT /tag= b
FT /note= "given in the specification but does not seem to
FT be part of the coding sequence and does not encode any
FT corresponding amino acids"
FT misc_feature 135..146
FT /tag= c

```

FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 207. .216
 FT /tag= d
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 277. .290
 FT /tag= e
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 351. .361
 FT /tag= f
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 422. .436
 FT /tag= g
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 497. .511
 FT /tag= h
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 572. .583
 FT /tag= i
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 644. .653
 FT /tag= j
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 714. .723
 FT /tag= k
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 784. .801
 FT /tag= l
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT
 FT WO200029428-A2.
 FT
 FT 25-MAY-2000.
 FT
 FT 18-NOV-1999; 99WO-GB003859.
 FT
 FT 18-NOV-1998; 98GB-00025303.
 FT 27-JAN-1999; 99GB-00001739.
 FT 30-JUL-1999; 99GB-00017995.
 FT
 FT (OXFO-) OXFORD BIOMEDICA UK LTD.
 FT
 FT Carroll MW, Myers KA;
 FT WPI; 2000-387735/33.
 FT P-PSDB; RAY94351.
 FT
 FT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 FT response useful in vaccinating against and in treating tumors.
 FT
 FT Disclosure; Page 78-79; 79pp; English.
 FT
 FT The present sequence encodes the canine 5T4 tumour-associated antigen
 FT (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in

CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC human or murine 5T4 gene sequence. The 5T4 antigen was shown to be
 CC effective at eliciting an immunotherapeutic anti-tumour response. Both
 CC the nucleic acid encoding the antigen and the antigen itself can be used
 CC to elicit an immune response, preferably CTL or an antibody response in a
 CC subject
 XX SQ Sequence 901 BP; 178 A; 246 C; 212 G; 153 T; 0 U; 112 Other;
 SQ
 Alignment Scores:
 Pred. No.: 93.6 Length: 901
 Score: 41.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 87.2% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-13 (1-9) x AAA27060 (1-901)
 QY 2 LeuTyLeuProArgAspValLeu 9
 Db 147 CTCTACTTGCCTCGCGACGTCCTG 170
 RESULT 34
 ADT16562/c
 ID ADT16562 standard; cDNA; 1268 BP.
 AC
 XX ADT16562;
 DT 13-JAN-2005 (first entry)
 XX
 DE Plant cDNA, Seq ID 1888.
 KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
 KW drought tolerance; disease resistance; galactomannan production;
 KW plant growth regulator; heat tolerance; herbicide tolerance;
 KW lignin production; extreme osmotic condition tolerance;
 KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
 KW seed protein yield.
 XX
 OS Viridiplantae.
 XX
 XX US2004216190-A1.
 XX
 PD 28-OCT-2004.
 XX
 XX 18-DEC-2003; 2003US-00739930.
 XX
 XX 28-APR-2003; 2003US-00424599.
 PR 28-APR-2003; 2003US-00425115.
 XX
 XX (KOVA/) KOVALIC D K.
 XX
 XX Kovalic DK;
 XX
 XX WPI; 2004-757369/74.
 XX
 XX New recombinant DNA constructs useful in the field of biochemistry and
 PT Genetics, and in particular for producing transgenic plants with improved
 PT biological characteristics.
 XX
 XX Claim 1; SEQ ID NO 1888; 14pp; English.
 XX
 CC The invention relates a recombinant DNA construct comprising a
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
 CC Arabidopsis, wheat and rape but the specification does not indicate which
 CC sequences is derived from which organism. Also included is a method of


```

PA (ASAH ) ASahi KASEI KK.
PI Sugahara T, Matsuuda A, Honda G, Muramatsu S, Ishizawa K;
DR WPI; 2004-122214/12.
DR P-PSDB; ADI26161.
XX
XX New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX
XX Claim 4; SEQ ID NO 125; 1368pp; English.
XX
XX The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infections disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.
XX
SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 508 Length: 2557
Score: 40.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 85.1% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-13 (1-9) x ADI26160 (1-2557)
QY 1 PheLeuTyrLeuProArgAspValLeu 9
Db 1237 TTCTTTTCCTCGCTCGGACTTACTA 1263
RESULT 37
ADI26158
ID ADI26158 standard; cDNA; 2557 BP.
XX
XX ADI26158;
AC
XX
XX 22-APR-2004 (first entry)
DT
XX
XX Human cDNA encoding protein that promotes STAT6 activation #62.
XX
XX ss; gene; human; signal transducer and activator of transcription 6;
KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
XX
XX Homo sapiens.
OS
XX
XX WO2003104277-A2.
PN

```

```

XX
PD
XX
XX 18-DEC-2003.
XX
XX 05-JUN-2003; 2003WO-JP007123.
XX
XX 05-JUN-2002; 2003JP-00164257.
PR
XX 06-JUN-2002; 2003US-0385912P.
PR
XX 26-DEC-2002; 2003JP-00377326.
PR
XX 27-DEC-2002; 2003US-0436467P.
PR
XX 15-MAY-2003; 2003JP-00137505.
PR
XX 16-MAY-2003; 2003US-0470836P.
XX
XX (ASAH ) ASahi KASEI KK.
PA
XX
XX Sugahara T, Matsuuda A, Honda G, Muramatsu S, Ishizawa K;
PI
XX WPI; 2004-122214/12.
DR
XX P-PSDB; ADI26159.
DR
XX
XX New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX
XX Claim 4; SEQ ID NO 123; 1368pp; English.
XX
XX The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infections disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.
XX
SQ Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 508 Length: 2557
Score: 40.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 85.1% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-13 (1-9) x ADI26158 (1-2557)
QY 1 PheLeuTyrLeuProArgAspValLeu 9
Db 1237 TTCTTTTCCTCGCTCGGACTTACTA 1263
RESULT 38
ADI25579/c
ID ADT45579 standard; cDNA; 2568 BP.
XX
XX ADT45579;
AC
XX
XX 02-DEC-2004 (first entry)
DT
XX

```

DE Bacterial polynucleotide #20330.

XX Recombinant DNA construct; transformed plant; improved plant property;

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KW bacterial polynucleotide; gene; ss.

XX Bacteria.

XX US2003233675-A1.

PN 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

PI WPI; 2004-061375/06.

DR New recombinant DNA construct comprising a promoter positioned to provide

XX for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 44017; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition. Improved lignin production or improved galactomannan

CC production. This sequence represents a bacterial polynucleotide used in

CC the scope of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2568 BP; 774 A; 484 C; 620 G; 690 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	510	Length:	2568
Score:	40.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	87.5%	Mismatches:	0
Query Match:	85.1%	Indels:	0
DB:	13	Gaps:	0

US-10-774-176-13 (1-9) x AD745579 (1-2568)

Qy 1 PheLeuTyrlLeuProArgAspVal 8

Db 219 TTCTTTATCTTCCTAAAGATGTT 196

RESULT 39

ID ADS46521/c

XX ADS46521 standard; cDNA; 2571 BP.

AC ADS46521;

XX 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #1264.

XX Recombinant DNA construct; transformed plant; improved plant property;

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KW bacterial polynucleotide; gene; ss.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

PI WPI; 2004-061375/06.

DR New recombinant DNA construct comprising a promoter positioned to provide

XX for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 24951; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition. Improved lignin production or improved galactomannan

CC production. This sequence represents a bacterial polynucleotide used in

CC the scope of the invention. Note: The sequence data for this patent did

CC not form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 2571 BP; 775 A; 484 C; 621 G; 691 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 511 Length: 2571
Score: 40.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 85.1% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-13 (1-9) x ADS46521 (1-2571)

QY 1 PheLeuTyrLeuProArgAspVal 8
DB 219 TTCCTTTATCTTCCTAAAGATGTT 196

RESULT 40

AAF97854

ID AAF97854 standard; DNA; 34488 BP.

XX
AC AAF97854;

DT 31-MAY-2001 (first entry)

XX Human neuroblastoma cell line NB-1 Ip36 nucleotide sequence SEQ ID NO:68.

XX Human; chromosome 1; Ip36; neuroblastoma cell line; NB-1; anticancer;
KW tumour suppressor; human Ip36 homozygosity deletion domain; tumour;
KW diagnosis; ds.

XX Homo sapiens.

XX WO200116311-A1.

XX 08-MAR-2001.

XX 31-AUG-2000; 2000WO-JP005930.

XX 31-AUG-1999; 99JP-00245962.

XX 09-MAY-2000; 2000JP-00136266.

XX (HISM) HISAMITSU PHARM CO LTD.
PA (CHIB-) CHIBA PREFECTURE.

XX Nakagawara A;

XX WPI; 2001-226686/23.

XX Human Ip36 homozygosity deletion domain from the 36-position of first
PT chromosome short arm in human neuroblastoma cell lines, applicable e.g.
PT in gene diagnosis of tumors as well as in developing anti-cancer drugs.

XX Example 8; Page 104-118; 226pp; Japanese.

XX The present invention describes a homozygosity deletion domain co-
CC existing in the 36-position of the first chromosome short arm (Ip36) in
CC human neuroblastoma. Also described are base sequences from the Ip36
CC position of human neuroblastoma cell lines (NB-1 and MASS-NB-SCH-1),
CC which are tumour suppressor genes in human neuroblastoma. The genes are
CC tumour suppressor genes, base sequence data of which are applicable as
CC tumour markers and reagents in studying mechanism of tumour body
CC formation, and gene diagnosis of tumours as well as in developing anti-
CC cancer drugs. AAF97787 to AAF97829 represent PCR primers used in the
CC exemplification of the present invention, and AAF97830 to AAF97874
CC represent sequences given in the exemplification of the present invention

XX Sequence 34488 BP; 9654 A; 6717 C; 6926 G; 11191 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.03e+04 Length: 34488
Score: 40.00 Matches: 6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0

Query Match: 85.1% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-13 (1-9) x AAF97854 (1-34488)

QY 1 PheLeuTyrLeuProArgAspValLeu 9

DB 17251 TTTTGTACGTACCAAGGATATTCTT 17277

RESULT 41

AAC66548

ID AAC66548 standard; DNA; 121162 BP.

XX
AC AAC66548;

XX 19-FEB-2001 (first entry)

XX Human kinesin-like protein HKLP coding sequence contig SEQ ID NO: 1.

XX Human; kinesin-like protein; HKLP; KIF1; cell division; cancer;
KW intracellular transport; neurological disorder; infertility;
KW biallelic marker; spontaneous abortion; neonatal chromosome disorder;
KW aneuploidy; ds.

XX Homo sapiens.

XX WO200063375-A1.

XX 26-OCT-2000.

XX 20-APR-2000; 2000WO-IB000562.

XX 20-APR-1999; 99US-0130217P.

XX (GEST) GENSET.

XX Bougueleret L, Dufaure-Gare I, Grel P;
PI WPI; 2000-665242/64.

XX An isolated or purified human kinesin-like protein (HKLP) encoding
PT polynucleotide used to detect HKLP polynucleotides in a sample comprises
PT a contiguous span of at least 12 nucleotides.

XX Claim 1; Page 143-175; 199pp; English.

XX The present invention describes the coding and protein sequences of the
CC human kinesin-like protein HKLP. It is thought that the protein could be
CC involved in neurological disorders, infertility, spontaneous abortion,
CC neonatal chromosome disorders, aneuploidy and cancers. This is due to its
CC function in the movement of microtubules. The protein shows homology to
CC the murine KIF1A and KIF1B proteins. The sequences disclosed in the
CC invention can be used in the isolation of similar human proteins and in
CC vector production. In addition, the biallelic markers shown can be used
CC in disease diagnosis and population studies

XX Sequence 121162 BP; 33272 A; 24108 C; 25842 G; 37919 T; 0 U; 21 Other;

Alignment Scores:

Pred. No.: 4.44e+04 Length: 121162
Score: 40.00 Matches: 6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match: 85.1% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-13 (1-9) x AAC66548 (1-121162)

QY 1 PheLeuTyrLeuProArgAspValLeu 9

DB 44101 TTTTGTACGTACCAAGGATATTCTT 44127

RESULT 42

AAQ34023/C
 ID AAQ34023 standard; DNA; 108 BP.
 AC AAQ34023;
 XX
 XX 25-MAR-2003 (revised)
 DT 02-FEB-1993 (first entry)
 XX
 XX Sequence upstream from a microsatellite from clone TGLA423.
 DE
 XX PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
 KW genetic mapping; traits; amplification; ss.
 XX
 OS Bos taurus.
 XX
 XX WO9213102-A1.
 PN
 XX
 XX 06-AUG-1992.
 PD
 XX
 XX 15-JAN-1992; 92WO-US000340.
 PF
 XX
 XX 15-JAN-1991; 91US-00642342.
 PR
 XX
 XX (GENM-) GENMARK.
 PA
 XX
 XX Georges M, Massey JM;
 PI
 XX
 XX WPI; 1992-284684/34.
 DR
 XX
 XX Polymorphic bovine DNA markers - used in genetic identification, gene
 PT mapping, and selective breeding.
 FT
 XX
 XX Table 7; Page 339; 517pp; English.
 PS
 XX
 XX The sequence is upstream of a bovine microsatellite sequence obtd. by
 CC screening a library of bovine MboI DNA fragments of between 250 and 500
 CC bp with an (AC)15 and a (TC)15 oligonucleotide probe. One out of 50
 CC clones cross-hybridised. Assuming independent distribution of
 CC microsatellites and MboI sites, the frequency of (T6)n > 9 microsatellites
 CC in the bovine genome is estimated at >100, 000. The sequence information
 CC for ca. 230 such bovine microsatellites is summarised in the
 CC specification and indexed herein (see below). The sequences upstream and
 CC downstream of the microsatellite sequence were used to generate the
 CC required PCR primers for in vitro amplification of the corresp.
 CC microsatellite (using the program OPTIPRIM). The microsatellites may be
 CC used to identify individuals, for parentage testing, and in the genetic
 CC mapping of economic trait loci, or genes involved in the determination of
 CC economically important traits esp. in cattle, to allow selective
 CC breeding. See also AAQ33501-34437. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX
 XX Sequence 108 BP; 37 A; 13 C; 18 G; 40 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 21 Length: 108
 Score: 39.00 Matches: 6
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 75.0% Mismatches: 0
 Query Match: 83.0% Indels: 0
 DB: 2 Gaps: 0
 US-10-774-176-13 (1-9) x AAQ34023 (1-108)
 QY 1 PheLeuTyrlEuProArgAspVal 8
 Db 80 TTTCATATCTACCCAAAGATATA 57
 RESULT 43
 ADQ49285/C
 ID ADQ49285 standard; DNA; 584 BP.
 XX
 AC ADQ49285;
 XX

21-OCT-2004 (first entry)
 Novel canine microarray-related DNA sequence SeqID587.
 canine microarray; drug screening; toxicity assay;
 environmental pollutant; cellular response; gene expression profile;
 toxic response; liver necrosis; fatty liver disease;
 protein adduct formation; hepatitis; dog; da.
 Canis familiaris.
 WO2004063324-A2.
 29-JUL-2004.
 05-MAY-2003; 2003WO-US013853.
 03-MAY-2002; 2002US-0377240P.
 (GENE-) GENE LOGIC INC.
 (PFIZ) PFIZER PROD INC.
 Diggins JC, Porter M, Wei T;
 WPI; 2004-561890/54.
 New isolated nucleic acid molecule, useful for drug screening and
 toxicity assays or for assessing the impact, including toxicity, of a
 compound, pharmaceutical agent or environmental pollutant on a cell or
 living organism.
 Claim 1; SEQ ID NO 587; 41pp; English.
 This invention is related to a novel isolated canine nucleic acid
 sequences and the construction of canine microarrays containing a
 significant portion of the canine genome. The isolated canine nucleic
 acid sequences of the invention may be useful for drug screening and
 toxicity assays. The invention is therefore useful for assessing the
 impact, including toxicity, of a compound, pharmaceutical agent or
 environmental pollutant on a cell or living organism. The methods are
 useful for detecting genes that are up- or down-regulated in canines in a
 disease state. The sequences are useful as diagnostic agents or markers
 to detect a cellular response in a sample individually or as part of a
 gene expression profile. It is also useful as a target for agents that
 modulate gene expression or activity. The database is useful for
 producing electronic Northern blots that allow the user to determine the cell
 type or tissue in which a given gene is expressed and to allow the
 determination of the abundance or expression level of a given gene in a
 particular tissue or cell. The methods are useful for determining the
 similarity of a toxic response to one or more individual compounds. The
 methods are useful for predicting at least one toxic response or the
 likelihood that a compound or test agent will induce various specific
 pathologies such as those of the liver (liver necrosis, fatty liver
 disease, protein adduct formation or hepatitis), those of the kidney,
 heart, brain or testes, or other pathologies associated with at least one
 of the toxins. The methods are also useful for predicting or elucidating
 the potential cellular pathways influenced, induced or modulated by the
 compound or test agent due to the similarity of the expression profile
 compared to the profile induced by a known toxin. The present sequence is
 that of a canine DNA sequence which was claimed for use during the
 production of a canine microarray of the invention.
 Sequence 584 BP; 149 A; 107 C; 127 G; 167 T; 0 U; 34 Other;
 Alignment Scores:
 Pred. No.: 149 Length: 584
 Score: 39.00 Matches: 6
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 75.0% Mismatches: 0
 Query Match: 83.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-13 (1-9) x ADQ49285 (1-584)

QY 1 PheLeuTyrLeuProArgAspVal 8
 DB 540 TATTGTATTACCGAGGACATT 517
 RESULT 44
 AEB53920
 ID AEB53920 standard; cDNA; 753 BP.
 AC AEB53920;
 XX
 XX
 DT 06-OCT-2005 (first entry)
 DE cDNA encoding prostate cancer-specific protein, seqid 67.
 XX screening; diagnosis; cancer; prostate tumor; cytostatic; drug screening;
 KW antibody therapy; immunoconjugate; immuno-diagnosis; immunotherapy;
 KW anion transport protein; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO2005067629-A2.
 XX
 XX 28-JUL-2005.
 XX
 XX 04-JAN-2005; 2005WO-US000040.
 XX
 XX 06-JAN-2004; 2004US-0534419P.
 XX
 XX (AVAL-) AVALON PHARM.
 XX
 XX Weigle B, Ebner R;
 XX WPI; 2005-563837/57.
 XX
 XX Identifying agent that modulates activity of cancer-related gene, by
 PT contacting compound with cell, to promote expression of gene, and
 PT detecting difference in expression of gene relative to when compound is
 PT not present, to identify agent.
 XX
 XX Claim 1; SEQ ID NO 67; 131pp; English.
 XX
 CC The present invention relates to methods of screening cancer-linked genes
 CC and expression products for cancer diagnosis, and for screening potential
 CC anti-cancer agents. Specifically claimed is a method of identifying (M1)
 CC an agent that modulates activity of cancer-related genes, by contacting a
 CC compound with a cell expressing a gene under conditions promoting the
 CC expression of gene, and detecting difference in expression relative to
 CC when the compound is not present. Also claimed are methods of identifying
 CC (M2) an anti-neoplastic agent, by contacting a cell exhibiting neoplastic
 CC activity with a compound first identified as a cancer related gene
 CC modulator by (M1), and detecting a decrease in the neoplastic, or
 CC administering agent identified by (M1) to an animal exhibiting a cancer
 CC condition and detecting a decrease in the cancerous condition. Also given
 CC is a protein (I), which has at least one immunogenic fragment; an
 CC antibody that reacts with a protein; an immunoconjugate comprising the
 CC antibody and a cytotoxic agent. The agent is useful for treating cancer,
 CC by contacting a cancerous cell in vivo with an agent having activity
 CC against an expression produced encoded by a gene sequence given in the
 CC specification. The agent is also useful for treating cancer in an animal,
 CC by eliciting the production of cytotoxic T lymphocytes specific for the
 CC protein. Also described is a method of diagnosing prostate cancer. The
 CC proteins are anion transport proteins specific for prostate cancer. The
 CC present sequence is cDNA encoding a prostate cancer-specific hypothetical
 CC protein CGI-96, unigene cluster Hs.360940.
 XX
 SQ Sequence 753 BP; 184 A; 194 C; 205 G; 170 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 199 Length: 753
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0

Query Match: 83.0% Indels: 0
 DB: 14 Gaps: 0
 US-10-774-176-13 (1-9) x AEB53920 (1-753)
 QY 1 PheLeuTyrLeuProArgAsp 7
 DB 159 TTTCGTACCTTCCCGAGAT 179
 RESULT 45
 ACC62259
 ID ACC62259 standard; cDNA; 937 BP.
 XX
 XX ACC62259;
 XX
 XX 23-JUN-2003 (first entry)
 DT Human NOV4b encoding cDNA SEQ ID NO:47.
 DE
 XX
 XX Human; NOVX; antiatherosclerotic; hypotensive; cardiac; dermatological;
 KW anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility;
 KW haemostatic; antiinflammatory; antiasthmatic; anti-HIV; immunomodulator;
 KW neuroprotective; nootropic; antiparkinsonian; metabolic; antilipemic;
 KW gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
 KW congenital heart defect; aortic stenosis; valve disease; transplantation;
 KW tuberosus sclerosis; obesity; congenital adrenal hyperplasia; diabetes;
 KW prostate cancer; metabolic disorder; neoplasm; lymphoma; uterus cancer;
 KW fertility; haemophilia; hypercoagulation; graft versus host disease;
 KW idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia;
 KW Crohn's disease; multiple sclerosis; infectious disease; cancer;
 KW cancer-associated cachexia; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia;
 KW metabolic syndrome X; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO2003023001-A2.
 PN
 XX 20-MAR-2003.
 PD
 XX
 XX 09-SEP-2002; 2002WO-US028538.
 PF
 XX 07-SEP-2001; 2001US-0318120P.
 PR 07-SEP-2001; 2001US-0318184P.
 PR 10-SEP-2001; 2001US-0318430P.
 PR 17-SEP-2001; 2001US-0322636P.
 PR 17-SEP-2001; 2001US-0322781P.
 PR 17-SEP-2001; 2001US-0322816P.
 PR 17-SEP-2001; 2001US-0322817P.
 PR 19-SEP-2001; 2001US-0323519P.
 PR 20-SEP-2001; 2001US-0323631P.
 PR 20-SEP-2001; 2001US-0323636P.
 PR 25-SEP-2001; 2001US-0324969P.
 PR 25-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324990P.
 PR 14-DEC-2001; 2001US-0341144P.
 PR 26-FEB-2002; 2002US-0359599P.
 PR 05-MAR-2002; 2002US-0361663P.
 PR 03-MAY-2002; 2002US-0377908P.
 PR 17-MAY-2002; 2002US-0381483P.
 PR 29-MAY-2002; 2002US-0383863P.
 PR 02-JUL-2002; 2002US-0393332P.
 PR 17-JUL-2002; 2002US-0396412P.
 PR 13-AUG-2002; 2002US-0403517P.
 PR 06-SEP-2002; 2002US-00236417.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Agee ML, Alsobrook JP, Anderson DM, Berghs C, Boldog FL;
 PI Burgess CE, Casman SJ, Catterton E, Chant JS, Chaudhuri A;
 PI Crabtree J, Dipippo VA, Edinger SR, Eisen AJ, Ellerman K;
 PI Gangoli EA, Gerlach VL, Giot L, Gorman L, Guo X, Gusev VY, Ji W;
 PI Kekuda R, Khrantsov NV, Leach MD, Lepley DM, Li L, Liu X;

PI Malyankar UM, Miller CE, Ooi CE, Ort T, Padigaru M, Patturajan M;
 PI Pena CEA, Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Voss EZ;
 PI Spaderna SK, Spytek KA, Taupier RJ, Twomlow N, Vernat CAM, Voss EZ;
 XX Zerhusen BD, Zhong M;
 DR WPI; 2003-313241/30.
 DR P-PSDB; ABR54190.

XX Novel human proteins and nucleic acid encoding the proteins, useful for
 PT diagnosis, treatment and prevention of disorders involving the human
 PT protein or nucleic acid e.g. cardiac and neurological disorders.

XX Claim 20; Page 118-119; 460pp; English.

XX The present invention describes isolated human NOVX proteins, where X is
 CC 1 to 42. ACC62236 to ACC62345 encodes the human NOVX proteins given in
 CC ABR54167 to ABR54276. NOVX sequences have antiatherosclerotic, cardiac,
 CC hypotensive, dermatological, anorectic, immunosuppressive, cytostatic,
 CC antidiabetic, antiinfertility, haemostatic, antiinflammatory, anti-HIV,
 CC antiasthmatic, metabolic, immunomodulator, neuroprotective, nootropic,
 CC antiparkinsonian and antilipemic activities, and can be used in gene
 CC therapy. NOVX proteins are useful for treating or preventing a pathology
 CC associated with a NOVX protein in humans and for treating a syndrome
 CC associated with the human disease. NOVX nucleic acids, proteins and
 CC antibodies can be used in the treatment and diagnosis of cardiomyopathy,
 CC atherosclerosis, hypertension, congenital heart defects, aortic stenosis,
 CC valve disease, tuberosclerosis, scleroderma, obesity, transplantation,
 CC congenital adrenal hyperplasia, prostate cancer, diabetes, metabolic
 CC disorders, neoplasm, lymphoma, uterus cancer, fertility, haemophilia,
 CC hypercoagulation, idiopathic thrombocytopenic purpura, graft versus host
 CC disease, AIDS, bronchial asthma, Crohn's disease, multiple sclerosis,
 CC infectious disease, anorexia, cancer-associated cachexia, cancer,
 CC Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, dylipidaemias, and metabolic syndrome X.
 CC ACC62346 to ACC62465 represent PCR primers and probes for human NOVX
 CC sequences, which are used in examples from the present invention.
 CC ABR54277 represents a human tryptophan protein given in comparison with
 CC the human NOV35b protein in the exemplification of the present invention

XX SQ Sequence 937 BP; 217 A; 257 C; 258 G; 205 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 257 Length: 937
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 83.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x ACC62259 (1-937)

QY 1 PheLeuTyrlEuProArGAsp 7
 Db 322 TTTCGTACCTTCCCGAGAT 342

RESULT 46

AAL57532

ID AAL57532 standard; cDNA; 945 BP.

XX AAL57532;

XX 06-NOV-2003 (first entry)

XX cDNA encoding human homologue of fruit fly kraken.

XX Anorectic; antiinflammatory; cardiac; hypotensive; antidiabetic;
 XX neuroprotective; pharmaceutical composition; body-weight regulation;
 KW thermogenesis; metabolic; obesity; Syndrome X; insulin-resistance;
 KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstone;
 KW pancreatic dysfunction; arteriosclerosis; coronary heart disease; gene;
 KW hypercholesterolaemia; dylipidaemia; osteoarthritis; ROS defence; ss;
 KW reactive oxygen species; neurodegenerative; mitochondrial; gene therapy;
 KW human; kraken.

XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 1..945
 FT /*tag= a
 FT /product= "Human homologue of fruit fly kraken protein"

XX WO2003061681-A2.

XX 31-JUL-2003.

XX 24-JAN-2003; 2003WO-EP000738.

XX 25-JAN-2002; 2002EP-00001806.

XX 14-FEB-2002; 2002EP-00003473.

XX 28-FEB-2002; 2002EP-00004687.

XX 25-APR-2002; 2002EP-00009475.

XX 18-JUN-2002; 2002EP-00013329.

XX 30-DEC-2002; 2002EP-00029081.

XX (DEVE-) DEVELOGEN ENTWICKLUNGSBIOLOGISCHE FORSCH.

XX Steuernagel A, Molitor A, Eulenberg K, Broemner G;

XX WPI; 2003-627418/59.

XX P-PSDB; AAO23978.

XX New pharmaceutical composition, useful for the manufacture of an agent
 PT for diagnosing, treating or preventing disorders related to body-weight
 PT regulation and thermogenesis, e.g., metabolic diseases such as obesity.
 XX Claim 2; Fig 17B; 144pp; English.

XX The invention relates to a novel pharmaceutical composition comprising a
 CC nucleic acid molecule or polypeptide which is a human homologue of a
 CC drosophila melanogaster polypeptide or polynucleotide. The composition of
 CC the invention may be utilised during the diagnosis, study, prevention and
 CC treatment of diseases related to body-weight regulation and thermogenesis
 CC including metabolic disorders such as obesity, Syndrome X and insulin-
 CC resistance syndrome and eating disorders e.g. cachexia, diabetes
 CC mellitus, hypertension, pancreatic dysfunctions, arteriosclerosis,
 CC coronary heart disease, hypercholesterolaemia, dylipidaemia,
 CC osteoarthritis and gallstones. Furthermore, disorders related to reactive
 CC oxygen species (ROS) defence may be addressed by the invention including
 CC neurodegenerative disorders or mitochondrial disorders. Finally, the
 CC composition of the invention may be useful in gene therapy. The current
 CC sequence is that of the cDNA encoding the human homologue of fruit fly
 CC kraken of the invention

XX SQ Sequence 945 BP; 217 A; 261 C; 254 G; 213 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 259 Length: 945
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 83.0% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-13 (1-9) x AAL57532 (1-945)

QY 1 PheLeuTyrlEuProArGAsp 7

Db 351 TTTCGTACCTTCCCGAGAT 371

RESULT 47

ADQ87521

ID ADQ87521 standard; cDNA; 945 BP.

XX ADQ87521;

XX 07-OCT-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #4399.
 XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KW cancer; cell proliferative disorder; gene; ss.
 XX Homo sapiens.
 XX WO2004060270-A2.
 XX 22-JUL-2004.
 XX 15-OCT-2003; 2003WO-US029126.
 XX 18-OCT-2002; 2002US-0418988P.
 XX (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 XX Wu TD, Zhou Y;
 XX WPI; 2004-534300/51.
 XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX Claim 1; SEQ ID NO 4399; 5504pp; English.
 XX The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.

SQ Sequence 945 BP; 217 A; 261 C; 254 G; 213 T; 0 U; 0 Other;

Alignment Scores: 259 Length: 945
 Pred. No.: 39.00 Matches: 7
 Score:

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 83.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-13 (1-9) x ADQ87521 (1-945)
 QY 1 PheLeuTyrLeuProArgAsp 7
 Db 351 TTTCTGTACCTTCCCGAGAT 371
 RESULT 48
 ACC62260
 ID ACC62260 standard; cDNA; 1058 BP.
 XX
 AC ACC62260;
 XX
 DT 23-JUN-2003 (first entry)
 XX
 DE Human NOV4c encoding cDNA SEQ ID NO:49.
 XX
 KW Human; NOVX; antiatherosclerotic; hypotensive; cardiant; dermatological;
 KW anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility;
 KW haemostatic; antiinflammatory; antiasthmatic; anti-HIV; immunomodulator;
 KW neuroprotective; nootropic; antiparkinsonian; metabolic; antilipemic;
 KW gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
 KW congenital heart defect; aortic stenosis; valve disease; transplantation;
 KW tuberosus sclerosis; obesity; congenital adrenal hyperplasia; diabetes;
 KW prostate cancer; metabolic disorder; neoplasm; lymphoma; uterus cancer;
 KW fertility; haemophilia; hypercoagulation; graft versus host disease;
 KW idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia;
 KW Crohn's disease; multiple sclerosis; infectious disease; cancer;
 KW cancer-associated cachexia; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia;
 KW metabolic syndrome X; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003023001-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 09-SEP-2002; 2002WO-US028538.
 XX
 PR 07-SEP-2001; 2001US-0318120P.
 PR 07-SEP-2001; 2001US-0318184P.
 PR 10-SEP-2001; 2001US-0318430P.
 PR 17-SEP-2001; 2001US-0322636P.
 PR 17-SEP-2001; 2001US-0322781P.
 PR 17-SEP-2001; 2001US-0322816P.
 PR 17-SEP-2001; 2001US-0322817P.
 PR 19-SEP-2001; 2001US-0323519P.
 PR 20-SEP-2001; 2001US-0323631P.
 PR 20-SEP-2001; 2001US-0323636P.
 PR 25-SEP-2001; 2001US-0324969P.
 PR 25-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324990P.
 PR 14-DEC-2001; 2001US-0341144P.
 PR 26-FEB-2002; 2002US-0359599P.
 PR 05-MAR-2002; 2002US-0361663P.
 PR 03-MAY-2002; 2002US-0377908P.
 PR 17-MAY-2002; 2002US-0381483P.
 PR 29-MAY-2002; 2002US-0383863P.
 PR 02-JUL-2002; 2002US-0393332P.
 PR 17-JUL-2002; 2002US-0396412P.
 PR 13-AUG-2002; 2002US-0403517P.
 PR 06-SEP-2002; 2002US-00236417.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;
 PI Burgess CE, Casman SJ, Catterton E, Chant JS, Chaudhuri A;
 PI Crabtree J, Dipippo VA, Edinger SR, Eise AJ, Ellerman K;

PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Gusev VY, Ji W;
 PI Kekuda R, Khramsov NV, Leach MD, Lepley DM, Li L, Liu X;
 PI Malyankar UM, Miller CE, Ort T, Padigara M, Patturajan M;
 PI Pena CEA, Rieger DK, Rothenberg ME, Shenoy SG, Shinkets RA;
 PI Spaderna SK, Spytek KA, Taupier RJ, Twomlow N, Vernet CAM, Voss EZ;
 PI Zerhusen BD, Zhong M;
 XX WPI; 2003-313241/30.
 DR P-PSDB; ABR54191.

XX Novel human proteins and nucleic acid encoding the proteins, useful for
 PT diagnosis, treatment and prevention of disorders involving the human
 PT protein or nucleic acid e.g. cardiac and neurological disorders.

XX Claim 20; Page 119; 460pp; English.

XX The present invention describes isolated human NOVX proteins, where X is
 CC 1 to 42. ACC62236 to ACC62345 encode the human NOVX proteins given in
 CC ABR54167 to ABR54276. NOVX sequences have antiatherosclerotic, cardiac,
 CC hypotensive, dermatological, anorectic, immunosuppressive, cytostatic,
 CC antidiabetic, antiinfertility, haemostatic, antiinflammatory, anti-HIV,
 CC antiaesthetic, metabolic, immunomodulator, neuroprotective, nootropic,
 CC antiparkinsonian and antilipase activities, and can be used in gene
 CC therapy. NOVX proteins are useful for treating or preventing a pathology
 CC associated with a NOVX protein in humans and for treating a syndrome
 CC associated with the human disease. NOVX nucleic acids, proteins and
 CC antibodies can be used in the treatment and diagnosis of cardiomyopathy,
 CC atherosclerosis, hypertension, congenital heart defects, aortic stenosis,
 CC valve disease, tuberosus sclerosis, scleroderma, obesity, transplantation,
 CC congenital adrenal hyperplasia, prostate cancer, diabetes, metabolic
 CC disorders, neoplasm, lymphoma, uterus cancer, fertility, haemophilia,
 CC hypercoagulation, idiopathic thrombocytopenic purpura, graft versus host
 CC disease, AIDS, bronchial asthma, Crohn's disease, multiple sclerosis,
 CC infectious disease, anorexia, cancer-associated cachexia, cancer,
 CC Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, dyslipidaemias, and metabolic syndrome X.
 CC ACC62346 to ACC62465 represent PCR primers and probes for human NOVX
 CC sequences, which are used in examples from the present invention.
 CC ABR54277 represents a human trypsinogen protein given in comparison with
 CC the human NOV35b protein in the exemplification of the present invention
 XX
 SQ Sequence 1058 BP; 243 A; 295 C; 284 G; 236 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 296 Length: 1058
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 83.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x ACC62260 (1-1058)

Qy 1 PheLeuTyrLeuProArgAsp 7
 Db 364 TTTCTGTACCTTCCCGAGAT 384

RESULT 49
 ABR53915
 ID ABR53915 standard; DNA; 1165 BP.
 XX
 AC ABR53915;
 XX
 DT 06-OCT-2005 (first entry)
 XX
 DE DNA encoding prostate cancer-specific protein, seqid 62.

XX screening; diagnosis; cancer; prostate tumor; cytostatic; drug screening;
 KW antibody therapy; immunoconjugate; immuno-diagnosis; immunotherapy;
 KW anion transport protein; ds; gene; splice variant.
 XX Homo sapiens.

PN WO2005067629-A2.

XX 28-JUL-2005.

XX 04-JAN-2005; 2005WO-US000040.

XX 06-JAN-2004; 2004US-0534419P.

XX (AVAL-) AVALON PHARM.

XX Weigle B, Ebner R;

XX WPI; 2005-563837/57.

DR P-PSDB; ABR53916.

XX Identifying agent that modulates activity of cancer-related gene, by
 PT contacting compound with cell, to promote expression of gene, and
 PT detecting difference in expression of gene relative to when compound is
 PT not present, to identify agent.

XX Claim 1; SEQ ID NO 62; 131pp; English.

XX The present invention relates to methods of screening cancer-linked genes
 CC and expression products for cancer diagnosis, and for screening potential
 CC anti-cancer agents. Specifically claimed is a method of identifying (M1)
 CC an agent that modulates activity of cancer-related genes, by contacting a
 CC compound with a cell expressing a gene under conditions promoting the
 CC expression of gene, and detecting difference in expression relative to
 CC when the compound is not present. Also claimed are methods of identifying
 CC (M2) an anti-neoplastic agent, by contacting a cell exhibiting neoplastic
 CC activity with a compound first identified as a cancer related gene
 CC modulator by (M1), and detecting a decrease in the neoplastic, or
 CC administering agent identified by (M1) to an animal exhibiting a cancer
 CC condition and detecting a decrease in the cancerous condition. Also given
 CC is a protein (I), which has at least one immunogenic fragment; an
 CC antibody that reacts with a protein; an immunoconjugate comprising the
 CC antibody and a cytotoxic agent. The agent is useful for treating cancer,
 CC by contacting a cancerous cell in vivo with an agent having activity
 CC against an expression product encoded by a gene sequence given in the
 CC specification. The agent is also useful for treating cancer in an animal,
 CC by eliciting the production of cytotoxic T lymphocytes specific for the
 CC protein. Also described is a method of diagnosing prostate cancer. The
 CC proteins are anion transport proteins specific for prostate cancer. The
 CC present sequence is variant 4 DNA encoding a prostate cancer-specific
 CC protein, unigene cluster Hs.360940.

XX SQ Sequence 1165 BP; 273 A; 319 C; 325 G; 248 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 331 Length: 1165
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 83.0% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-13 (1-9) x ABR53915 (1-1165)

Qy 1 PheLeuTyrLeuProArgAsp 7
 Db 419 TTTCTGTACCTTCCCGAGAT 439

RESULT 50
 ADI40459
 ID ADI40459 standard; DNA; 1241 BP.
 XX
 AC ADI40459;
 XX
 DT 22-APR-2004 (first entry)
 XX Human purified secretory polynucleotide (sptm), seq id 152.
 DE
 XX Antiarteriosclerotic; antipsoriatic; cytostatic; secretory molecule;

KW agonist; antagonist; gene therapy; antisense; human; secretory;
KW purified secretory polynucleotide; sptm; toxicity; arteriosclerosis;
KW psoriasis; cancer; gene; da.

XX Homo sapiens.

XX WO2003062385-A2.

XX 31-JUL-2003.

XX 15-JAN-2003; 2003WO-US001605.

XX 17-JAN-2002; 2002US-0349413P.

XX 17-JAN-2002; 2002US-0349946P.

XX (INCY-) INCYTE GENOMICS INC.

PA (JONE/) JONES A L.

PA (DAHL/) DAHL C R.

PA (GIET/) GIETZEN D.

PA (CHIN/) CHINN J.

PA (DUFO/) DUFOUR G E.

PA (JACK/) JACKSON J L.

PA (YUJY/) YU J Y.

PA (TUAS/) TUASON O.

PA (YAPP/) YAP P E.

PA (AMSH/) AMSHEY S R.

PA (DAMT/) DAM T C.

PA (LIUT/) LIU T F.

PA (GERS/) GERSTIN B H.

PA (PERA/) PERALTA C H.

PA (LEWI/) LEWIS S A.

PA (CHEN/) CHEN A J.

PA (MARW/) MARWAHA R.

PA (LANR/) LAN R Y.

PA (URAS/) URASHKA M E.

PA (KRIS/) KRISTNAM S R.

PA (KOLL/) KOLLURU V.

PA (PANE/) PANESAR I S.

XX Jones AL, Dahl CR, Gietzen D, Chinn J, Dufour GE, Jackson JL;

PI Yu JY, Tuason O, Yap PS, Amshey SR, Dam TC, Liu TF, Gerstin EH;

PI Peralta CH, Lewis SA, Chen AJ, Marwaha R, Lan RY, Urashka ME;

PI Kristnam SR, Kolluru V, Panesar IS;

XX WPI: 2003-853444/79.

DR P-PSDB; ADI40622.

XX New isolated secreted polynucleotide for diagnosing or treating

PT conditions, diseases or disorders associated with cell signaling e.g.

PT arteriosclerosis, psoriasis, and cancer.

XX Claim 1a; SEQ ID NO 152; 486pp; English.

XX The invention relates to isolated purified secreted polynucleotides
CC (sptm) (I), and the polypeptides (SPTM) encoded by sptm. A polypeptide
CC encoded by (I) is used to identify a compound which binds to it. A
CC microarray comprising (I) is used to generate a transcript. (I) is used
CC to screen a compound for effectiveness in altering expression of (I). (I)
CC is used to assess toxicity of a test compound. An agonist or antagonist
CC identified by a new method is used in a pharmaceutical composition. The
CC secretory molecules are used to diagnose or treat conditions, diseases or
CC disorders associated with cell signaling e.g. arteriosclerosis,
CC psoriasis, and cancer. Sequences given in ADI40308-ADI40468 represent
CC human purified secretory polynucleotides of the invention, and those
CC given in ADI40469-ADI40631 represent the polypeptides they encode.

XX SQ Sequence 1241 BP; 321 A; 291 C; 281 G; 348 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	356	Length:	1241
Score:	39.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0

Query Match: 83.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-13 (1-9) x ADI40459 (1-1241)

QY 1 PheLeuTyrLeuProArgAsp 7

DB 128 TTCTCTATTATTACCAAGAGAT 148

Search completed: April 25, 2006, 12:32:33

Job time : 341.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using (xframe plug p2n model)

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds

(without alignments)

171.290 Million cell updates/sec

Title: US-10-774-176-13

Perfect score: 47

Sequence: 1 FLYLPDVL 9

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/abes/ABSWEB_spool/US10774176/runat_24042006_165114_19197/app.query.fasta_1
-DB=GenEmbl -Qfmt=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowsum2 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR score=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04
-USER=US10774176 @CGN_1_1_6765 @runat_24042006_165114_19197 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	47	100.0	927	6	AX829164 Sequence
2	47	100.0	1263	6	BD249731 Polypepti
3	47	100.0	1263	6	AX025011 Sequence

4	47	100.0	1263	6	AX149553 Sequence
5	47	100.0	1263	6	AX316086 Sequence
6	47	100.0	1263	6	AX467371 Sequence
7	47	100.0	2053	8	CQ731678 Sequence
8	47	100.0	2053	8	HST40A
9	47	100.0	2359	6	BD127282 Primer fo
10	47	100.0	2359	6	CQ782724 Sequence
11	47	100.0	2359	8	AK074786 Homo sapi
12	47	100.0	2361	6	BD127283 Primer fo
13	47	100.0	2361	6	CQ782726 Sequence
14	47	100.0	2361	6	AX961916 Sequence
15	47	100.0	2361	8	AK074790 Homo sapi
16	47	100.0	2379	8	BC037161 Homo sapi
17	47	100.0	2714	8	AB168308 Macaca fa
18	47	100.0	5551	8	AB168308 Macaca fa
19	47	100.0	121909	8	HSJ492P14
20	46	97.9	28770	8	AF527803 Homo sapi
21	46	97.9	34669	8	AC000048 Homo sapi
22	46	97.9	101155	8	AL449423 Human DNA
23	46	97.9	250000	8	AB060808 Homo sapi
24	45	95.7	192971	8	AC040173 Homo sapi
25	45	95.7	194551	8	AC092291 Homo sapi
26	44	93.6	2333	9	AF063939 Rattus no
27	44	93.6	2361	9	BC087011 Rattus no
28	44	93.6	210237	14	AC128294 Rattus no
29	44	93.6	239076	14	AC106962 Rattus no
30	43	91.5	1260	6	AX467373 Sequence
31	43	91.5	1260	6	AX821533 Sequence
32	43	91.5	1260	6	AX821548 Sequence
33	43	91.5	26599	2	U40187 Caenorhabdi
34	43	91.5	27297	1	AY275838 Escherich
35	43	91.5	86248	6	BD184766 Nucleic a
36	43	91.5	87563	6	AR204161 Sequence
37	43	91.5	87563	6	AR637510 Sequence
38	43	91.5	110000	1	AE005174_11
39	43	91.5	110000	1	AE005174_15
40	43	91.5	110000	1	BA000007_14
41	43	91.5	110000	14	AC091229_04
42	43	91.5	110000	14	AC091242_4
43	43	91.5	110000	14	AC091347_4
44	43	91.5	134580	14	AC025359 Homo sapi
45	43	91.5	157393	9	AC124199 Mus muscu
46	43	91.5	184940	14	AL672264 Mus muscu
47	43	91.5	200087	8	AL354821 Human DNA
48	43	91.5	210180	14	AC110466 Rattus no
49	43	91.5	243114	9	AL672244 Mouse DNA
50	43	91.5	24964	14	AC137320 Rattus no
51	43	91.5	260560	14	AC137180 Rattus no
52	42	89.4	93381	14	AL18504 Homo sapi
53	42	89.4	95483	14	HSDJ766D4
54	42	89.4	140367	9	AC163908 Bos tauru
55	42	89.4	165097	8	AL732439 Mouse DNA
56	42	89.4	167013	8	AF003060 Homo sapi
57	42	89.4	180316	14	HSDJ828H9
58	41	87.2	901	6	BD249733 Polypepti
59	41	87.2	901	6	AX025013 Sequence
60	41	87.2	901	6	AX316088 Sequence
61	41	87.2	1702	5	CR848128 Xenopus t
62	41	87.2	18158	5	AY277972 Takifugu
63	41	87.2	45494	14	AC002318 Homo sapi
64	41	87.2	74760	14	AC166847 Bos tauru
65	41	87.2	110000	15	AP008215_111
66	41	87.2	110000	15	AP008215_112
67	41	87.2	138807	8	AL161650 Human DNA
68	41	87.2	140406	5	AC145796 Xenopus t
69	41	87.2	150001	8	AC006063
70	41	87.2	160410	15	AP005685 Oryza sat
71	41	87.2	172286	14	AC115258 Rattus no
72	41	87.2	180751	14	AC159354 Bos tauru
73	41	87.2	186313	15	AP005689 Oryza sat
74	41	87.2	192728	14	AC156957 Bos tauru
75	41	87.2	197000	2	TBBCHRL1
76	41	87.2	199113	8	AL445645 Human DNA

77	41	87.2	208409	14	AC117895	AC117895 Rattus no	c 150	40	85.1	202693	14	AC149784	AC149784 Bos tauru
78	41	87.2	210223	9	AC122183	AC122183 Mus muscu	151	40	85.1	203223	14	AC108997	AC108997 Rattus no
c 79	41	87.2	220807	14	AC121311	AC121311 Mus muscu	152	40	85.1	203689	14	AC163565	AC163565 Bos tauru
80	41	87.2	221011	14	AC128761	AC128761 Rattus no	c 153	40	85.1	206002	9	AC132470	AC132470 Mus muscu
81	41	87.2	222043	14	AC114154	AC114154 Rattus no	c 154	40	85.1	208587	14	AC079498	AC079498 Mus muscu
c 82	41	87.2	229073	9	AC134339	AC134339 Mus muscu	c 155	40	85.1	209834	9	AC147513	AC147513 Mus muscu
c 83	41	87.2	251966	14	AC164305	AC164305 Mus muscu	156	40	85.1	214228	14	AC153180	AC153180 Bos tauru
84	41	87.2	262031	14	AC115274	AC115274 Rattus no	c 157	40	85.1	214511	14	AC132556	AC132556 Rattus no
c 85	41	87.2	275945	14	AC112068	AC112068 Rattus no	c 158	40	85.1	214825	5	CR626874	CR626874 Zebrafish
86	41	87.2	276589	14	AC114039	AC114039 Rattus no	c 159	40	85.1	215167	14	AC153558	AC153558 Mus muscu
c 87	41	87.2	276624	14	AC132560	AC132560 Rattus no	c 160	40	85.1	224290	14	AC098199	AC098199 Rattus no
c 88	40	85.1	323	13	AF185890	AF185890 HIV-1 LP9	161	40	85.1	236697	14	AC127845	AC127845 Rattus no
c 89	40	85.1	447	10	BV312322	BV312322 S2365608	c 162	40	85.1	238226	14	CR925739	CR925739 Danio rer
90	40	85.1	668	10	BV376842	BV376842 S231613R	c 163	40	85.1	240932	9	AC098206	AC098206 Rattus no
c 91	40	85.1	795	3	AY711730	AY711730 Unculture	c 164	40	85.1	247652	9	AC156555	AC156555 Mus muscu
c 92	40	85.1	1060	5	BX932209	BX932209 Gallus ga	c 165	40	85.1	250101	14	AC106644	AC106644 Rattus no
93	40	85.1	1281	6	BD249732	BD249732 Polypepti	c 166	40	85.1	258910	14	AC129287	AC129287 Rattus no
94	40	85.1	1281	6	AX025012	AX025012 Sequence	167	40	85.1	261538	14	AC112004	AC112004 Rattus no
95	40	85.1	1281	6	AX316087	AX316087 Sequence	c 168	40	85.1	262637	14	AC135394	AC135394 Rattus no
96	40	85.1	2423	9	BC058198	BC058198 Mus muscu	169	40	85.1	269572	14	AC108552	AC108552 Rattus no
97	40	85.1	2557	6	AX961912	AX961912 Sequence	c 170	40	85.1	276534	14	AC149757	AC149757 Bos tauru
98	40	85.1	2557	6	AX961914	AX961914 Sequence	171	40	85.1	312314	14	AC160478	AC160478 Bos tauru
99	40	85.1	7942	9	MMU012160	MMU012160 Mus muscu	c 172	40	85.1	328866	14	AC107203	AC107203 Homo sapi
100	40	85.1	34488	6	BD093713	BD093713 Human lp3	c 173	39	83.0	494	10	G53414	G53414 SHGC-86244
101	40	85.1	50283	5	BX511091	BX511091 Zebrafish	c 174	39	83.0	601	6	AR662756	AR662756 Sequence
102	40	85.1	50833	8	AC079607	AC079607 Homo sapi	c 175	39	83.0	601	6	AR662757	AR662757 Sequence
c 103	40	85.1	50988	8	AC104776	AC104776 Homo sapi	c 176	39	83.0	601	6	AR662758	AR662758 Sequence
104	40	85.1	59008	14	AC100379	AC100379 Mus muscu	c 177	39	83.0	601	6	AR666647	AR666647 Sequence
105	40	85.1	73388	8	AC135036	AC135036 Homo sapi	c 178	39	83.0	601	6	AR666648	AR666648 Sequence
c 106	40	85.1	87358	8	AL137878	AL137878 Human DNA	c 179	39	83.0	601	6	AR666649	AR666649 Sequence
107	40	85.1	90289	14	AP003164	AP003164 Homo sapi	180	39	83.0	618	10	BV234203	BV234203 S234P6244
c 108	40	85.1	108031	5	BX511068	BX511068 Zebrafish	181	39	83.0	627	10	BV039963	BV039963 S212P6126
c 109	40	85.1	110000	1	BA000001_07	Continuation (8 of	182	39	83.0	684	8	BC093888	BC093888 Homo sapi
c 110	40	85.1	110000	1	BA000012_54	Continuation (55 o	183	39	83.0	753	8	H822813E	H822813E Novel hum
c 111	40	85.1	110000	1	BA000012_55	Continuation (56 o	c 184	39	83.0	757	10	BV486565	BV486565 S215P6034
c 112	40	85.1	110000	14	AC026388_0	AC026388 Mus muscu	c 185	39	83.0	822	10	BV594922	BV594922 S215P6165
113	40	85.1	110000	15	AP008209_317	Continuation (318	c 186	39	83.0	854	10	BV466791	BV466791 G531P6425
c 114	40	85.1	121162	6	AX039602	AX039602 Sequence	187	39	83.0	855	10	BV570448	BV570448 G591P6568
c 115	40	85.1	126871	14	AC163381	AC163381 Xenopus t	188	39	83.0	945	6	AX960249	AX960249 Sequence
116	40	85.1	129096	14	AC160903	AC160903 Loxodonta	189	39	83.0	945	8	H822813C	H822813C Novel hum
117	40	85.1	132841	5	BX004761	BX004761 Zebrafish	c 190	39	83.0	1041	2	AY274385	AY274385 Drosophil
c 118	40	85.1	136353	14	AC160891	AC160891 Loxodonta	c 191	39	83.0	1041	2	AY274386	AY274386 Drosophil
c 119	40	85.1	143583	8	HS467D16	AL009031 Human DNA	192	39	83.0	1346	8	AL365513	AL365513 Novel hum
120	40	85.1	147542	5	BX511234	BX511234 Zebrafish	193	39	83.0	1374	8	HS2228131	HS2228131 Novel hum
c 121	40	85.1	148750	8	AC002326	AC002326 Genomic s	c 194	39	83.0	1386	6	CQ491185	CQ491185 Sequence
c 122	40	85.1	148758	14	CR932437	CR932437 Danio rer	c 195	39	83.0	1386	6	CQ497050	CQ497050 Sequence
c 123	40	85.1	148996	14	AC155446	AC155446 Zea mays	196	39	83.0	1431	6	AX829180	AX829180 Sequence
c 124	40	85.1	150743	15	AC133450	AC133450 Oryza sat	c 197	39	83.0	1712	6	AX833269	AX833269 Sequence
c 125	40	85.1	154273	8	AL390857	AL390857 Human DNA	c 198	39	83.0	1712	8	AX094872	AX094872 Homo sapi
c 126	40	85.1	155220	9	AC151476	AC151476 Mus muscu	c 199	39	83.0	1989	6	BD127123	BD127123 Primer fo
c 127	40	85.1	155775	14	AC011943	AC011943 Homo sapi	c 200	39	83.0	1989	6	CQ782420	CQ782420 Sequence
c 128	40	85.1	158366	14	BX546449	BX546449 Danio rer	c 201	39	83.0	1989	8	AK074492	AK074492 Homo sapi
c 129	40	85.1	158906	14	AC097003	AC097003 Rattus no	c 202	39	83.0	2000	6	AX508662	AX508662 Sequence
c 130	40	85.1	159197	14	AC166187	AC166187 Xenopus t	c 203	39	83.0	2112	8	BC041802	BC041802 Homo sapi
c 131	40	85.1	159413	14	AC155128	AC155128 Bos tauru	c 204	39	83.0	2374	6	AX835254	AX835254 Sequence
c 132	40	85.1	160904	14	AC067896	AC067896 Homo sapi	c 205	39	83.0	2374	8	AK098201	AK098201 Homo sapi
c 133	40	85.1	161521	9	AC123810	AC123810 Mus muscu	c 206	39	83.0	2375	15	AF220204	AF220204 Malus dom
c 134	40	85.1	162951	14	AC084094	AC084094 Homo sapi	c 207	39	83.0	2678	6	CQ491584	CQ491584 Sequence
c 135	40	85.1	167046	9	AC158516	AC158516 Mus muscu	c 208	39	83.0	2678	6	CQ497471	CQ497471 Sequence
c 136	40	85.1	167998	14	AC026382	AC026382 Mus muscu	c 209	39	83.0	3953	6	CQ580853	CQ580853 Sequence
c 137	40	85.1	168085	14	AC096867	AC096867 Rattus no	c 210	39	83.0	9640	1	SARPOABC	X14818 Sulfolobus
c 138	40	85.1	171054	9	AC116590	AC116590 Mus muscu	c 211	39	83.0	12250	9	AF463754	AF463754 Mus muscu
c 139	40	85.1	171541	5	CR352298	CR352298 Zebrafish	c 212	39	83.0	18042	8	AL590794	AL590794 Human DNA
140	40	85.1	171612	8	AL358013	AL358013 Human DNA	c 213	39	83.0	37106	14	AC163737	AC163737 Pan trogl
c 141	40	85.1	175462	5	BX469912	BX469912 Zebrafish	c 214	39	83.0	37897	9	AC090444	AC090444 Rattus no
c 142	40	85.1	176043	14	AC113393	AC113393 Homo sapi	c 215	39	83.0	42064	2	CBRG4A014	CBRG4A014 Caenorhab
c 143	40	85.1	178000	9	AC107707	AC107707 Mus muscu	c 216	39	83.0	43417	14	BX294390	BX294390 Homo sapi
c 144	40	85.1	178925	14	AC161245	AC161245 Mus muscu	c 217	39	83.0	47665	15	AB022211	AB022211 Arabidops
c 145	40	85.1	189056	14	CT025607	CT025607 Mus muscu	c 218	39	83.0	48860	8	AC127390	AC127390 Homo sapi
c 146	40	85.1	189221	14	AC162501	AC162501 Bos tauru	c 219	39	83.0	55068	6	AR658650	AR658650 Sequence
c 147	40	85.1	190088	9	AC132446	AC132446 Mus muscu	c 220	39	83.0	64575	9	AC087213	AC087213 Rattus no
c 148	40	85.1	197530	9	AC099174	AC099174 Rattus no	c 221	39	83.0	64575	9	AC087213	AC087213 Rattus no
c 149	40	85.1	197655	9	AC123230	AC123230 Mus muscu	c 222	39	83.0	68407	14	AC048372	AC048372 Homo sapi

223	39	83.0	72949	14	AC069357	AC069357 Homo sapi	296	39	83.0	185748	14	AC137308	AC137308 Rattus no
224	39	83.0	81648	4	AC097351	AC097351 Sus scrofa	297	39	83.0	186127	8	AC104582	AC104582 Homo sapi
225	39	83.0	82309	8	AC138136	AC138136 Homo sapi	298	39	83.0	187072	8	AC099541	AC099541 Homo sapi
226	39	83.0	90724	8	AL159161	AL159161 Human DNA	299	39	83.0	187914	14	AC141124	AC141124 Rattus no
227	39	83.0	95107	8	AL591242	AL591242 Human DNA	c 300	39	83.0	188514	9	AC132374	AC132374 Mus muscu
228	39	83.0	98300	8	AC083790	AC083790 Homo sapi	c 301	39	83.0	188522	9	AC149587	AC149587 Mus muscu
229	39	83.0	100000	8	AP000074	AP000074 Homo sapi	c 302	39	83.0	191540	9	AC125203	AC125203 Mus muscu
230	39	83.0	101216	14	AC060227	AC060227 Homo sapi	303	39	83.0	192097	9	AC141878	AC141878 Mus muscu
231	39	83.0	106711	8	AL365205	AL365205 Human DNA	304	39	83.0	192169	8	AC099560	AC099560 Homo sapi
232	39	83.0	108315	8	HS41P2	AL049748 Human DNA	c 305	39	83.0	193008	14	AC026676	AC026676 Homo sapi
233	39	83.0	110000	1	CP000177_05	Continuation (6 of	c 306	39	83.0	193098	8	AL137140	AL137140 Human DNA
234	39	83.0	110000	3	AP004180_2	Continuation (3 of	307	39	83.0	193860	9	BL119959	BL119959 Mouse DNA
235	39	83.0	110000	14	AP006498_3	Continuation (4 of	c 308	39	83.0	194624	8	AC008742	AC008742 Homo sapi
236	39	83.0	110000	14	CT009548_2	Continuation (3 of	c 309	39	83.0	194880	14	AC091929	AC091929 Homo sapi
237	39	83.0	112303	8	AC093836	AC093836 Homo sapi	310	39	83.0	196119	9	AC161410	AC161410 Mus muscu
238	39	83.0	115863	8	HSJ537P22	AL109733 Human DNA	311	39	83.0	196311	9	AC159644	AC159644 Mus muscu
239	39	83.0	119361	14	AC092771	AC092771 Homo sapi	312	39	83.0	196388	9	AC153867	AC153867 Mus muscu
240	39	83.0	119430	15	ATP9D16	AL035394 Arabidops	c 313	39	83.0	197633	9	AC140301	AC140301 Mus muscu
241	39	83.0	120873	8	AL445986	AL445986 Human DNA	314	39	83.0	198433	9	AC127327	AC127327 Mus muscu
242	39	83.0	123360	8	HS22E2E13	Z93241 Human DNA s	315	39	83.0	199199	15	ATCHRIV59	AL161559 Arabidops
243	39	83.0	129467	8	HSJ601K24	AL109626 Human DNA	316	39	83.0	199334	14	AC133664	AC133664 Rattus no
244	39	83.0	129995	5	EX897665	EX897665 Zebrafish	317	39	83.0	199882	8	AL354720	AL354720 Human DNA
245	39	83.0	130039	8	AL355587	AL355587 Human DNA	c 318	39	83.0	201978	9	AC149285	AC149285 Mus muscu
246	39	83.0	133392	9	AC113633	AC113633 Rattus no	319	39	83.0	203634	14	CNS01DVL	AL135879 Homo sapi
247	39	83.0	133406	2	AF321227	AF321227 Tribolium	320	39	83.0	203650	8	CNS01DSI	AL121790 Human chr
248	39	83.0	139255	14	AC141580	AC141580 Rattus no	c 321	39	83.0	204809	8	AC021133	AC021133 Homo sapi
249	39	83.0	143669	8	AP006307	AP006307 Homo sapi	322	39	83.0	206044	9	AC138210	AC138210 Mus muscu
250	39	83.0	145070	14	AC132544	AC132544 Rattus no	323	39	83.0	206466	14	AC109663	AC109663 Rattus no
251	39	83.0	146385	8	AC010534	AC010534 Homo sapi	324	39	83.0	206840	14	AC147743	AC147743 Mus muscu
252	39	83.0	148018	14	AC068861	AC068861 Homo sapi	c 325	39	83.0	207841	14	AC072019	AC072019 Homo sapi
253	39	83.0	148497	14	AC079592	AC079592 Homo sapi	326	39	83.0	213099	14	AC132659	AC132659 Rattus no
254	39	83.0	149940	14	AC138560	AC138560 Lemur cat	327	39	83.0	216550	14	AC036233	AC036233 Homo sapi
255	39	83.0	151162	8	AC024621	AC024621 Homo sapi	328	39	83.0	217496	9	AC153630	AC153630 Mus muscu
256	39	83.0	151273	9	AC133190	AC133190 Mus muscu	329	39	83.0	218047	14	AC132168	AC132168 Rattus no
257	39	83.0	152202	8	AC094081	AC094081 Homo sapi	c 330	39	83.0	218889	9	AC079959	AC079959 Mus muscu
258	39	83.0	153783	14	AC163058	AC163058 Bos tauru	c 331	39	83.0	219557	14	AC162991	AC162991 Bos tauru
259	39	83.0	154557	14	AC135458	AC135458 Felis cat	c 332	39	83.0	223090	14	AC111554	AC111554 Rattus no
260	39	83.0	156212	8	HS126B4	AL022316 Human DNA	c 333	39	83.0	223491	9	AC129600	AC129600 Mus muscu
261	39	83.0	157301	2	AC159409	AC159409 Trypanoso	c 334	39	83.0	223734	14	AC073750	AC073750 Mus muscu
262	39	83.0	158341	8	AC021733	AC021733 Homo sapi	c 335	39	83.0	229546	9	AC123625	AC123625 Mus muscu
263	39	83.0	159500	9	AC005742	AC005742 Mus muscu	c 336	39	83.0	229797	14	AC133348	AC133348 Rattus no
264	39	83.0	160671	8	AC090451	AC090451 Homo sapi	337	39	83.0	231636	14	AC144444	AC144444 Rattus no
265	39	83.0	160725	14	HS919B11	AL031302 Homo sapi	338	39	83.0	233613	14	AC135537	AC135537 Rattus no
266	39	83.0	162401	14	AC015646	AC015646 Homo sapi	c 339	39	83.0	235438	14	AC129004	AC129004 Rattus no
267	39	83.0	163734	14	AC148134	AC148134 Atelerix	340	39	83.0	236005	14	CR627492	CR627492 Danio rer
268	39	83.0	164098	14	AC010756	AC010756 Homo sapi	341	39	83.0	236017	14	AC136909	AC136909 Rattus no
269	39	83.0	164221	14	AC148133	AC148133 Atelerix	342	39	83.0	237819	14	CR392369	CR392369 Danio rer
270	39	83.0	164293	14	AC166652	AC166652 Mus muscu	c 343	39	83.0	237823	14	AC020886	AC020886 Mus muscu
271	39	83.0	165270	14	AC015765	AC015765 Homo sapi	344	39	83.0	238675	14	AC095962	AC095962 Rattus no
272	39	83.0	165292	9	AC123049	AC123049 Mus muscu	c 345	39	83.0	238802	14	AC127450	AC127450 Rattus no
273	39	83.0	166452	8	AC090453	AC090453 Homo sapi	c 346	39	83.0	238980	14	AC106296	AC106296 Rattus no
274	39	83.0	166522	14	AL357072	AL357072 Homo sapi	347	39	83.0	239840	14	AC094367	AC094367 Rattus no
275	39	83.0	167569	8	AL160338	AL160338 Mus muscu	348	39	83.0	240414	14	AC110423	AC110423 Rattus no
276	39	83.0	167596	14	AC013177	AC013177 Drosophil	c 349	39	83.0	241285	14	AC157609	AC157609 Mus muscu
277	39	83.0	168242	8	AC020692	AC020692 Homo sapi	c 350	39	83.0	241783	14	AC106100	AC106100 Rattus no
278	39	83.0	168608	14	AL591477	AL591477 Homo sapi	c 351	39	83.0	242701	14	AC102972	AC102972 Rattus no
279	39	83.0	170609	8	AC113132	AC113132 Homo sapi	352	39	83.0	242907	14	AC106660	AC106660 Rattus no
280	39	83.0	170624	8	AC006031	AC006031 Homo sapi	353	39	83.0	246113	14	AC079581	AC079581 Mus muscu
281	39	83.0	170917	8	AC121761	AC121761 Homo sapi	354	39	83.0	246173	14	AC110096	AC110096 Rattus no
282	39	83.0	170943	9	AL596095	AL596095 Mouse DNA	c 355	39	83.0	246173	14	AC110096	AC110096 Rattus no
283	39	83.0	173268	14	AC112087	AC112087 Rattus no	356	39	83.0	247249	14	AC023526	AC023526 Homo sapi
284	39	83.0	173534	8	AC007920	AC007920 Homo sapi	357	39	83.0	247544	14	AC092741	AC092741 Mus muscu
285	39	83.0	175940	8	HSJ15217	AL109318 Human DNA	358	39	83.0	250722	14	AC159733	AC159733 Bos tauru
286	39	83.0	175994	8	AC092337	AC092337 Homo sapi	c 359	39	83.0	251556	14	AC152779	AC152779 Bos tauru
287	39	83.0	176267	14	AC153028	AC153028 Bos tauru	360	39	83.0	258835	14	AC114130	AC114130 Rattus no
288	39	83.0	176565	14	AL136106	AL136106 Homo sapi	361	39	83.0	262798	14	AC111702	AC111702 Rattus no
289	39	83.0	176748	14	AC111837	AC111837 Rattus no	362	39	83.0	266623	14	AC148859	AC148859 Ootlemur
290	39	83.0	179134	14	AC119607	AC119607 Rattus no	363	39	83.0	268049	14	AC129466	AC129466 Rattus no
291	39	83.0	181823	14	AC145526	AC145526 Papio ham	c 364	39	83.0	271507	2	AE003625	AE003625 Drosophil
292	39	83.0	183007	2	AC007257	AC007257 Drosophil	365	39	83.0	271848	14	AC094384	AC094384 Rattus no
293	39	83.0	183149	8	AC007909	AC007909 Homo sapi	c 366	39	83.0	274636	14	AC114077	AC114077 Rattus no
294	39	83.0	185297	14	AC148846	AC148846 Ootlemur	367	39	83.0	283037	14	AC158086	AC158086 Bos tauru
295	39	83.0	185698	14	AC119065	AC119065 Canis fam	368	39	83.0	294050	1	BX294142	BX294142 Pirellula

C 369	39	83.0	302446	14	AC120576	AC120576 Rattus no	C 442	38	80.9	114532	8	AL389883	AL389883 Human DNA
C 370	39	83.0	306749	14	AL590310	AL590310 Homo sapi	443	38	80.9	118312	5	CR759819	CR759819 Zebrafish
C 371	39	83.0	325916	14	AL531174	AL531174 Bos tauru	444	38	80.9	119301	5	CR388064	CR388064 Zebrafish
C 372	39	83.0	342586	14	AC109682	AC109682 Rattus no	445	38	80.9	119658	14	AP007523	AP007523 Lotus cor
C 373	39	83.0	343188	14	AC114853	AC114853 Rattus no	C 446	38	80.9	119696	8	HSDJ7719	AL049794 Human DNA
C 374	39	83.0	349635	14	AC093978	AC093978 Rattus no	447	38	80.9	120169	8	AC005060	AC005060 Homo sapi
C 375	38	80.9	231	10	AB139793	AB139793 Homo sapi	448	38	80.9	120364	15	AC140915	AC140915 Medicago
C 376	38	80.9	342	6	CQ710684	CQ710684 Sequence	449	38	80.9	126004	5	EX914219	EX914219 Zebrafish
C 377	38	80.9	378	10	AB146153	AB146153 Homo sapi	C 450	38	80.9	130872	9	AL929106	AL929106 Mouse DNA
C 378	38	80.9	515	6	BD120896	BD120896 EST and e	451	38	80.9	133379	15	AC083835	AC083835 Arabidops
C 379	38	80.9	515	6	AR425343	AR425343 Sequence	452	38	80.9	133980	8	AC124817	AC124817 Brachioles
C 380	38	80.9	515	6	AX986037	AX986037 Sequence	453	38	80.9	135684	8	AC002463	AC002463 Homo sapi
C 381	38	80.9	585	6	CQ919017	CQ919017 Sequence	454	38	80.9	136641	14	AC114327	AC114327 Canis fam
C 382	38	80.9	574	10	BV297955	BV297955 S239P694F	C 455	38	80.9	136641	14	AC114327	AC114327 Canis fam
C 383	38	80.9	588	10	BV303357	BV303357 S239P6328	456	38	80.9	138838	5	CR354606	CR354606 Zebrafish
C 384	38	80.9	698	6	AR612916	AR612916 Sequence	457	38	80.9	139791	8	AC137673	AC137673 Homo sapi
C 385	38	80.9	708	15	AY596779	AY596779 Lycopersi	458	38	80.9	142068	15	AP003723	AP003723 Oryza sat
C 386	38	80.9	720	15	AY949616	AY949616 Lycopersi	459	38	80.9	142514	5	EX548069	EX548069 Zebrafish
C 387	38	80.9	755	10	BV621845	BV621845 S215P6071	C 460	38	80.9	142773	14	AC153104	AC153104 Felis cat
C 388	38	80.9	996	5	CR387003	CR387003 Gallus ga	C 461	38	80.9	144136	14	AC012243	AC012243 Homo sapi
C 389	38	80.9	1230	5	CR386347	CR386347 Gallus ga	C 462	38	80.9	144400	9	AL731832	AL731832 Mouse DNA
C 390	38	80.9	1378	8	HSJ323804	AJ323804 Homo sapi	463	38	80.9	145115	14	AC092197	AC092197 Homo sapi
C 391	38	80.9	1777	5	AF044977	AF044977 Danio rer	C 464	38	80.9	146351	5	AL935207	AL935207 Zebrafish
C 392	38	80.9	1857	15	AK106660	AK106660 Oryza sat	465	38	80.9	149479	14	CR394568	CR394568 Danio rer
C 393	38	80.9	2561	5	AF077225	AF077225 Danio rer	C 466	38	80.9	149512	14	CR382281	CR382281 Danio rer
C 394	38	80.9	3222	5	AF013242	AF013242 Xenopus l	C 467	38	80.9	149726	14	AC021970	AC021970 Homo sapi
C 395	38	80.9	4438	8	HSM807017	BM641115 Homo sapi	468	38	80.9	150021	5	EX088691	EX088691 Zebrafish
C 396	38	80.9	6508	15	PVA005763	AJ005763 Phaseolus	C 469	38	80.9	151194	9	AL112988	AL112988 Mus muscu
C 397	38	80.9	11566	1	AE001637	AE001637 Chlamydia	470	38	80.9	151328	14	EX936340	EX936340 Danio rer
C 398	38	80.9	14353	1	AE002184	AE002184 Chlamydia	C 471	38	80.9	152715	8	AP003050	AP003050 Homo sapi
C 399	38	80.9	30494	2	CEK10D6	Z74040 Caenorhabdi	C 472	38	80.9	152794	8	AL139376	AL139376 Human DNA
C 400	38	80.9	32690	14	AY714841	AY714841 Unculture	C 473	38	80.9	153113	14	AC128832	AC128832 Rattus no
C 401	38	80.9	35984	14	AC140910	AC140910 Homo sapi	C 474	38	80.9	153448	8	AC063923	AC063923 Homo sapi
C 402	38	80.9	43276	8	AC074214	AC074214 Homo sapi	475	38	80.9	153562	8	AC009046	AC009046 Homo sapi
C 403	38	80.9	49255	14	AC009114	AC009114 Homo sapi	C 476	38	80.9	153801	14	CR387932	CR387932 Danio rer
C 404	38	80.9	55281	14	AC109922	AC109922 Arabidops	C 477	38	80.9	154297	14	CR628330	CR628330 Danio rer
C 405	38	80.9	57853	14	AC164967	AC164967 Bos tauru	C 478	38	80.9	154607	9	AL139868	AL139868 Mus muscu
C 406	38	80.9	60004	14	AC013601	AC013601 Homo sapi	C 479	38	80.9	154639	14	AC090175	AC090175 Homo sapi
C 407	38	80.9	64091	14	AC100990	AC100990 Mus muscu	480	38	80.9	155995	14	AC034162	AC034162 Homo sapi
C 408	38	80.9	65987	14	AC153871_3	Continuation (4 of	C 481	38	80.9	156339	14	AC068613	AC068613 Homo sapi
C 409	38	80.9	67550	14	AC069034	AC069034 Homo sapi	C 482	38	80.9	156415	8	AC011464	AC011464 Homo sapi
C 410	38	80.9	69024	5	CR749767	CR749767 Zebrafish	C 483	38	80.9	157336	14	AC079787	AC079787 Homo sapi
C 411	38	80.9	72932	14	AC015300	AC015300 Drosophill	484	38	80.9	158134	5	AL953878	AL953878 Zebrafish
C 412	38	80.9	79534	5	CR628324	CR628324 Zebrafish	485	38	80.9	159082	14	AC013649	AC013649 Homo sapi
C 413	38	80.9	85386	5	AC098704	AC098704 Danio rer	C 486	38	80.9	159225	5	EX28753	EX28753 Zebrafish
C 414	38	80.9	86476	15	NCB8B20	AL355933 Neurospor	C 487	38	80.9	159405	14	AC115908	AC115908 Mus muscu
C 415	38	80.9	90015	8	AL513264	AL513264 Human DNA	C 488	38	80.9	160395	14	AC152860	AC152860 Felis cat
C 416	38	80.9	92140	5	EX914207	EX914207 Zebrafish	C 489	38	80.9	161227	14	AC164546	AC164546 Mus muscu
C 417	38	80.9	94202	5	EX914220	EX914220 Zebrafish	490	38	80.9	161518	14	EX679660	EX679660 Danio rer
C 418	38	80.9	96694	8	AC090470	AC090470 Homo sapi	C 491	38	80.9	161523	8	AC012186	AC012186 Homo sapi
C 419	38	80.9	100598	8	AC006271	AC006271 Homo sapi	492	38	80.9	161603	5	AL953880	AL953880 Zebrafish
C 420	38	80.9	101080	14	EX004860	EX004860 Danio rer	C 493	38	80.9	161687	8	AC048351	AC048351 Homo sapi
C 421	38	80.9	101241	5	AL845510	AL845510 Zebrafish	494	38	80.9	161858	5	CR749168	CR749168 Zebrafish
C 422	38	80.9	104757	14	AC133066	AC133066 Danio rer	495	38	80.9	162546	5	EX548051	EX548051 Zebrafish
C 423	38	80.9	106216	8	AC061997	AC061997 Homo sapi	496	38	80.9	162689	14	AC117145	AC117145 Rattus no
C 424	38	80.9	109440	8	AC092891	AC092891 Homo sapi	497	38	80.9	162695	14	AC037466	AC037466 Homo sapi
C 425	38	80.9	110000	1	BA000001_11	Continuation (12 o	498	38	80.9	163093	5	EX682544	EX682544 Zebrafish
C 426	38	80.9	110000	1	BA000008_05	Continuation (6 of	499	38	80.9	163288	14	CR812925	CR812925 Danio rer
C 427	38	80.9	110000	1	BA000008_06	Continuation (7 of	500	38	80.9	164819	8	AC092335	AC092335 Homo sapi
C 428	38	80.9	110000	6	AR310754_06	Continuation (7 of	C 501	38	80.9	167722	8	AC073548	AC073548 Homo sapi
C 429	38	80.9	110000	6	AR607478_05	Continuation (6 of	C 502	38	80.9	168172	8	AC010528	AC010528 Homo sapi
C 430	38	80.9	110000	6	AR607478_06	Continuation (7 of	503	38	80.9	168222	5	AL929286	AL929286 Zebrafish
C 431	38	80.9	110000	14	AC139251_3	Continuation (4 of	C 504	38	80.9	168334	14	AC027208	AC027208 Homo sapi
C 432	38	80.9	110000	14	AC140833_3	Continuation (4 of	C 505	38	80.9	168370	9	AC147104	AC147104 Mus muscu
C 433	38	80.9	110000	14	AC141815_1	Continuation (2 of	506	38	80.9	168837	8	AC090108	AC090108 Homo sapi
C 434	38	80.9	110000	15	CR380954_02	Continuation (3 of	507	38	80.9	169004	14	AC121682	AC121682 Rattus no
C 435	38	80.9	110000	15	CR382130_21	Continuation (22 o	508	38	80.9	169032	8	AC016931	AC016931 Homo sapi
C 436	38	80.9	110000	15	CR382139_19	Continuation (20 o	C 509	38	80.9	169201	9	AC123676	AC123676 Mus muscu
C 437	38	80.9	110000	15	AE017351_09	Continuation (10 o	C 510	38	80.9	169400	14	AC022208	AC022208 Homo sapi
C 438	38	80.9	110000	15	AP008212_279	Continuation (280	511	38	80.9	170736	14	AC013658	AC013658 Homo sapi
C 439	38	80.9	110627	14	AP008031	AP008031 Lotus cor	512	38	80.9	171428	8	AP003062	AP003062 Homo sapi
C 440	38	80.9	111669	8	AC124293	AC124293 Homo sapi	513	38	80.9	171486	5	EX539346	EX539346 Zebrafish
C 441	38	80.9	113879	8	AC011904	AC011904 Homo sapi	514	38	80.9	172193	14	AC113567	AC113567 Canis fam

661	37	78.7	256	6	AX779019	Sequence	CS015841	Sequence
662	37	78.7	400	10	G13785	human STR S	U29382	Caenorhabdi
663	37	78.7	436	3	AY934373	Unculture	AL021106	Drosophil
664	37	78.7	551	10	G76785	S209P6214FG	AY177663	Homo sapi
665	37	78.7	585	6	Q0520296	Sequence	U88168	Caenorhabdi
666	37	78.7	590	10	BV226496	S233P6408	Z81533	Caenorhabdi
667	37	78.7	609	10	BV443703	S237P6337	AC000040	Homo sapi
668	37	78.7	621	10	BV412571	S229P6477	AC024762	Caenorhab
669	37	78.7	636	10	BV460745	gbp42f05.	AL512633	Human DNA
670	37	78.7	648	2	AY268381	Bombus di	AC145702	Homo sapi
671	37	78.7	650	10	G96514	S210P6083RG	AL390196	Human DNA
672	37	78.7	747	15	AK105369	Oryza sat	AL157918	S.pombe C
673	37	78.7	758	10	BV497094	S217P6104	AC137833	Homo sapi
674	37	78.7	759	6	BD150395	Primer fo	AF176815	Homo sapi
675	37	78.7	759	6	AX870333	Sequence	AC053476	Homo sapi
676	37	78.7	790	10	BV062581	S212P6754	AC026339	Homo sapi
677	37	78.7	818	6	Q0799317	Sequence	AC096343_3	Continuation (4 of
678	37	78.7	834	10	BV056065	S212P6220	AY357582	Burkholder
679	37	78.7	838	10	BV458290	ghf74b03.	AC100888	Mus muscu
680	37	78.7	1000	15	CN0311781	Cymodocea	AC090304	Homo sapi
681	37	78.7	1173	6	AX645205	Sequence	AC080165	Homo sapi
682	37	78.7	1173	6	AX656994	Sequence	AL499610	Human DNA
683	37	78.7	1244	8	BC074906	Homo sapi	AL591133	Human DNA
684	37	78.7	1244	8	BC074907	Homo sapi	AC080055	Homo sapi
685	37	78.7	1265	8	HUMPROT2	Human vitam	AL512637	Human DNA
686	37	78.7	1323	15	BT000069	Arabidops	AB006424	Bacillus
687	37	78.7	1415	15	MECGT1	X77459 M.esculenta	AC121763	Genomic B
688	37	78.7	1441	5	AJ721080	Gallus ga	BS000024	Pan trogl
689	37	78.7	1485	6	Q0722889	Sequence	AC100361	Mus muscu
690	37	78.7	1485	8	HUMPROX1	Human prote	AC101353	Mus muscu
691	37	78.7	1551	6	AX409708	Sequence	AC101581	Mus muscu
692	37	78.7	1551	8	HUMPROX2	Human prote	AL357059	Human DNA
693	37	78.7	1579	15	AY062860	Arabidops	AC165934	Bos tauru
694	37	78.7	1579	15	AY085962	Arabidops	AC016593	Bos tauru
695	37	78.7	1583	15	AY089125	Arabidops	AC010820	Homo sapi
696	37	78.7	1608	15	AF123390	Arabidops	AC008389	Homo sapi
697	37	78.7	1638	15	AB009567	Caenorhab	AC008389	Homo sapi
698	37	78.7	1714	1	AB213658	Vibrio ch	AC164826	Bos tauru
699	37	78.7	1751	6	AR204327	Sequence	AL356052	Human DNA
700	37	78.7	1751	6	AR637676	Sequence	AL731549	Human DNA
701	37	78.7	1761	6	BD184858	Nucleic a	AC015629	Homo sapi
702	37	78.7	1866	1	ECU82290	Escherichia	AC010950	Homo sapi
703	37	78.7	1916	15	AK104323	Oryza sat	AC019811	Drosophil
704	37	78.7	2169	15	AK067627	Oryza sat	AC011459	Homo sapi
705	37	78.7	2179	15	AK070388	Oryza sat	AC009324	Arabidops
706	37	78.7	2370	15	ZMA430205	Sequence	AC120337	Homo sapi
707	37	78.7	2412	9	BC016523	Mus muscu	AC135388	Homo sapi
708	37	78.7	2521	6	Q0608519	Sequence	AX928745	Danio rer
709	37	78.7	3118	6	AX780422	Sequence	AL136117	Human DNA
710	37	78.7	3118	8	D84430	Homo sapien	AC022513	Homo sapi
711	37	78.7	3232	8	AC000043	Homo sapi	AC06836	Arabidops
712	37	78.7	3262	1	AY136815	Spiroplas	AC068200	Homo sapi
713	37	78.7	3662	6	Q0608483	Sequence	AC141322	Medicago
714	37	78.7	4147	9	AK129004	Mus muscu	AC009185	Homo sapi
715	37	78.7	4179	9	BC007135	Mus muscu	AC068824	Homo sapi
716	37	78.7	4185	9	AB004817	Mus muscu	AC165547	Bos tauru
717	37	78.7	5667	6	BD157183	Primer fo	AP005472	Oryza sat
718	37	78.7	5667	6	AX878377	Sequence	AL670757	Mouse DNA
719	37	78.7	5667	8	AK002209	Homo sapi	AX695695	Sequence
720	37	78.7	6324	2	AY176645	Cherax qu	AC084402	Mus muscu
721	37	78.7	7008	15	AB018251	Magnaport	AC098580	Homo sapi
722	37	78.7	8240	8	HSU94396	Human dystr	AL929537	Mouse DNA
723	37	78.7	9077	1	AB032799	Chromobac	AL929537	Mouse DNA
724	37	78.7	10094	1	AF172851	Chromobac	AL024493	Human DNA
725	37	78.7	13097	1	AB000673	Aquifex a	AP006865	Lotus cor
726	37	78.7	13595	1	AE004397	Vibrio ch	AL139278	Human DNA
727	37	78.7	14871	8	HUMDYSTR0	Human dystr	AX695566	Sequence
728	37	78.7	16470	5	AL147909	Xenopus t	AC073101	Homo sapi
729	37	78.7	17581	9	AL591673	Mouse DNA	BX890635	Zebrafish
730	37	78.7	25869	8	AC002459	Homo sapi	AC151460	Medicago
731	37	78.7	27824	2	CER07B5	Caenorhabdi	AC166378	Bos tauru
732	37	78.7	29186	8	AL357560	Human DNA	AP007964	Lotus cor
733	37	78.7	32460	6	CS015837	Sequence	AC159298	Mus muscu

807	37	78.7	100180	15	AP004007	AP004007 Oryza sat	c 880	37	78.7	139254	9	AL672204	Mouse DNA
c 808	37	78.7	100731	14	AC138540	AC138540 Magnaport	881	37	78.7	130116	8	AL355572	Human DNA
809	37	78.7	101846	8	HS197L1	AL033390 Human DNA	882	37	78.7	130269	9	AC111067	Mus muscu
810	37	78.7	101944	8	AC011463	AC011463 Homo sapi	883	37	78.7	131747	14	CT009479	Medicago
811	37	78.7	106235	14	AC166443	AC166443 Bos tauru	884	37	78.7	133271	15	AC104277	Oryza sat
c 812	37	78.7	107870	15	AC105262	AC105262 Oryza sat	885	37	78.7	133791	8	AC091975	Homo sapi
c 813	37	78.7	108661	8	AC026736	AC026736 Homo sapi	886	37	78.7	134683	14	AC128075	Rattus no
814	37	78.7	110000	1	RE005174_49	Continuation (50 o	887	37	78.7	135580	14	AC017687	AC017687 Drosophil
c 815	37	78.7	110000	1	RS553306_31	Continuation (32 o	888	37	78.7	135959	15	AC144721	Medicago
c 816	37	78.7	110000	1	CR848038_06	Continuation (7 of	c 889	37	78.7	137165	14	AC011264	Homo sapi
817	37	78.7	110000	1	BA000007_48	Continuation (49 o	c 890	37	78.7	137302	4	AY386695	Trypanoso
c 818	37	78.7	110000	1	BA000028_13	Continuation (14 o	891	37	78.7	137759	2	TBN19B2	Human DNA
819	37	78.7	110000	1	BA000028_23	Continuation (24 o	c 892	37	78.7	138918	8	AL445433	Human DNA
820	37	78.7	110000	1	BA000028_24	Continuation (25 o	c 893	37	78.7	139446	4	AY386697	Oryctolag
c 821	37	78.7	110000	1	BA000045_18	Continuation (19 o	894	37	78.7	139599	14	AC141919	Rattus no
822	37	78.7	110000	14	AC091343_0	AC091343 Rattus no	895	37	78.7	139790	8	AL606827	Human DNA
c 823	37	78.7	110000	14	AC094411_2	Continuation (3 of	c 896	37	78.7	139803	5	EX247873	Zebrafish
824	37	78.7	110000	14	AC094411_3	Continuation (4 of	897	37	78.7	141424	5	AL845435	Zebrafish
c 825	37	78.7	110000	14	AC095948_2	Continuation (3 of	c 898	37	78.7	141921	9	AC130843	Mus muscu
826	37	78.7	110000	14	AC096343_2	Continuation (3 of	899	37	78.7	141955	8	AL158201	Human DNA
827	37	78.7	110000	14	AC105718_1	Continuation (2 of	c 900	37	78.7	142196	15	AC130728	Oryza sat
c 828	37	78.7	110000	14	AC111394_2	Continuation (3 of	901	37	78.7	142576	8	AC010609	Homo sapi
829	37	78.7	110000	14	AC140833_3	Continuation (4 of	902	37	78.7	142719	8	AC083803	Homo sapi
c 830	37	78.7	110000	14	AC146850_2	Continuation (3 of	c 903	37	78.7	143266	14	AC067790	Homo sapi
831	37	78.7	110000	14	AP006502_02	Continuation (3 of	c 904	37	78.7	143777	14	AC147817	Xenopus t
c 832	37	78.7	110000	14	EX284640_2	Continuation (3 of	c 905	37	78.7	144240	8	AC103703	Homo sapi
833	37	78.7	110000	14	EX322593_0	Continuation (3 of	c 906	37	78.7	144799	9	AC155804	Mus muscu
834	37	78.7	110000	14	CT009753_08	Continuation (9 of	c 907	37	78.7	145290	14	AC114242	Rattus no
835	37	78.7	110000	15	AP008214_230	Continuation (231	c 908	37	78.7	146438	8	HSBA191L9	Human DNA
836	37	78.7	110000	15	AP008214_230	Continuation (231	c 909	37	78.7	147054	14	AL356582	Homo sapi
c 837	37	78.7	110000	15	AP008214_230	Continuation (231	c 910	37	78.7	147242	14	AC135152	Rattus no
838	37	78.7	110000	15	CR382129_24	Continuation (25 o	c 911	37	78.7	147394	14	CR936414	Danio rer
c 839	37	78.7	110000	15	CR382133_01	Continuation (2 of	912	37	78.7	147444	14	AC027424	Homo sapi
840	37	78.7	110000	15	AE016820_01	Continuation (2 of	c 913	37	78.7	147895	14	AC140163	Homo sapi
c 841	37	78.7	110000	15	AE017342_01	Continuation (2 of	c 914	37	78.7	149705	15	AY286123	Cyanidios
842	37	78.7	110000	15	AP008211_293	Continuation (24 o	c 915	37	78.7	149790	8	AC022106	Homo sapi
c 843	37	78.7	110000	15	AP008211_293	Continuation (294	916	37	78.7	149858	14	AC159569	Callithri
844	37	78.7	110000	15	AP008212_111	Continuation (112	c 917	37	78.7	149879	14	AP005381	Oryza sat
c 845	37	78.7	110000	15	AP008212_175	Continuation (176	918	37	78.7	149987	15	AB002583	Cyanidios
846	37	78.7	110000	15	AP008212_281	Continuation (282	c 919	37	78.7	150172	8	AC006285	Homo sapi
c 847	37	78.7	110000	15	AP008213_223	Continuation (224	920	37	78.7	150236	14	CR936448	Danio rer
848	37	78.7	11081	14	AC010249	AC010249 Homo sapi	921	37	78.7	150601	8	AC117426	Homo sapi
c 849	37	78.7	111104	8	AC117389	AC117389 Homo sapi	922	37	78.7	150638	8	AP000874	Homo sapi
850	37	78.7	111590	8	HSN28L15	AL122015 Human DNA	c 923	37	78.7	150831	8	AC007719	Homo sapi
851	37	78.7	111764	8	AL133379	AL133379 Human DNA	c 924	37	78.7	150986	9	AL645587	Mouse DNA
852	37	78.7	111936	15	AC024260	AC024260 Arabidops	925	37	78.7	151018	8	AC073636	Homo sapi
c 853	37	78.7	112322	15	AC155228	AC155228 Medicago	926	37	78.7	151383	8	AC134980	Homo sapi
854	37	78.7	112831	8	HS287L14	Z95325 Human DNA s	c 927	37	78.7	151727	14	AC006894	Caenorhab
855	37	78.7	113329	2	AC073246	AC073246 Trypanoso	c 928	37	78.7	152369	14	CNS07ED2	Homo sapi
c 856	37	78.7	114428	8	AC092468	AC092468 Homo sapi	c 929	37	78.7	152537	8	AC055731	Homo sapi
857	37	78.7	115301	14	AC138541	AC138541 Magnaport	c 930	37	78.7	152572	8	AC090588	Homo sapi
c 858	37	78.7	115967	8	AC002112	AC002112 Genomic s	931	37	78.7	152881	14	AC025737	Homo sapi
859	37	78.7	116004	15	AC149208	AC149208 Medicago	932	37	78.7	153072	14	AC151525	Medicago
c 860	37	78.7	116457	8	AL137845	AL137845 Human DNA	933	37	78.7	153206	8	AC006356	Homo sapi
861	37	78.7	116498	5	AC149149	AC149149 Xenopus t	c 934	37	78.7	153700	9	AC118114	Rattus no
862	37	78.7	117886	14	AC139385	AC139385 Rattus no	c 935	37	78.7	154273	14	AC127577	Mus muscu
c 863	37	78.7	118030	14	AP000667	AP000667 Homo sapi	c 936	37	78.7	154297	5	EX072583	Zebrafish
864	37	78.7	118113	14	AC164318	AC164318 Bos tauru	c 937	37	78.7	154489	14	AC021069	Homo sapi
c 865	37	78.7	119915	8	AC133124	AC133124 Homo sapi	938	37	78.7	154814	8	AC093533	Homo sapi
866	37	78.7	120135	8	AC096506	AC096506 Homo sapi	c 939	37	78.7	155065	8	HSJ659119	Human DNA
c 867	37	78.7	121101	14	AC008034	AC008034 Homo sapi	c 940	37	78.7	155298	14	AC152845	Otolemur
868	37	78.7	122320	8	AC104058	AC104058 Homo sapi	941	37	78.7	156239	5	AC147211	Xenopus t
c 869	37	78.7	122427	8	AC025459	AC025459 Homo sapi	c 942	37	78.7	157007	14	AC066599	Homo sapi
870	37	78.7	123372	9	AC134551	AC134551 Mus muscu	c 943	37	78.7	157345	5	CNS098BU	Homo sapi
c 871	37	78.7	123570	8	AC022846	AC022846 Homo sapi	944	37	78.7	157872	8	AC007850	Homo sapi
872	37	78.7	123570	14	AC142034	AC142034 Rattus no	945	37	78.7	157879	8	AC149230	Fan trogi
c 873	37	78.7	124162	14	AC011121	AC011121 Rattus no	946	37	78.7	158134	14	AC135617	Homo sapi
874	37	78.7	125291	8	AP000646	AP000646 Homo sapi	c 947	37	78.7	158643	14	AC154144	Mus muscu
c 875	37	78.7	125445	8	AC078795	AC078795 Homo sapi	948	37	78.7	158656	5	EX511156	Zebrafish
876	37	78.7	126118	8	AC104598	AC104598 Homo sapi	c 949	37	78.7	159158	9	AC140463	Mus muscu
877	37	78.7	127561	8	AC114748	AC114748 Homo sapi	950	37	78.7	159178	5	EX248085	Zebrafish
c 878	37	78.7	128341	8	AC091675	AC091675 Homo sapi	951	37	78.7	159188	5	CR381658	Zebrafish
879	37	78.7	128639	5	EX548009	EX548009 Zebrafish	952	37	78.7	159512	8	AC133552	Homo sapi

AC126450 Mus muscu
AC025605 Homo sapi
AC150919 Bos tauru
284488 Human DNA s
AC125158 Mus muscu
AC079606 Trypanoso
AC068775 Homo sapi
AP004277 Oryza sat
AC146188 Pan trogl
AC079634 Genomic S
AC093903 Homo sapi
AC010139 Homo sapi
AC158842 Bos tauru
AC012186 Homo sapi
AP001929 Homo sapi
AC023947 Homo sapi
AC136667 Rattus no
AL138797 Homo sapi
AC068718 Homo sapi
AC131648 Homo sapi
AC023574 Homo sapi
AL731773 Mouse DNA
AC104460 Homo sapi
AC026799 Homo sapi
AC159449 Trypanoso
AL391475 Human DNA
AP004585 Oryza sat
AP004684 Oryza sat
AL355308 Homo sapi
BX571822 Zebrafish
AC158972 Mus muscu
AC019156 Homo sapi
AC153327 Callicebu
AL445533 Homo sapi
AC092874 Sus scrof
AC145442 Bos tauru
AC132311 Mus muscu
AL158142 Homo sapi
AC022677 Homo sapi
CT025564 Mus muscu
AL365448 Homo sapi
AC011271 Homo sapi
AC155717 Mus muscu
AC058815 Homo sapi
AC139120 Didelphis
AL929135 Zebrafish
AL391809 Human DNA

ALIGNMENTS

RESULT 1
AX829164
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .927
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.03 Length: 927
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x AX829164 (1-927)
Qy 1 PheLeuTyrLeuProArgAspValLeu 9
Db 322 TTCTTTTACTGCGCGGGATGTCTG 348
RESULT 2
BD249731
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
PD 17-SEP-2002
PF 18-NOV-1999 JP 200582415
PR 18-NOV-1998 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
PC C07K19/00,
PC C12N15/00,
CC Polypeptide
FH Key
FT source
Location/Qualifiers
1. .1263
/organism="Homo sapiens (human)"
FEATURES
source
1. .1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.47 Length: 1263
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x BD249731 (1-1263)
Qy 1 PheLeuTyrLeuProArgAspValLeu 9
Db 664 TTCTTTTACTGCGCGGGATGTCTG 690
RESULT 3
AX025011
LOCUS
AX025011 1263 bp DNA linear PAT 15-SEP-2000

```

RESULT 5
AX316086
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 1 from Patent EPI160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: Ep 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. .1263
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.47 Length: 1263
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x AX316086 (1-1263)

Oy 1 PheLeuTyrlauProArgAspValleu 9
|||||
Db 664 TTCCTTTACCTGCCGCGGATGCTG 690

RESULT 6
AX3167371
LOCUS AX3167371 1263 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX3167371
VERSION AX3167371.1 GI:21900602
KEYWORDS
SOURCE
ORGANISM Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. .1263
Location/Qualifiers
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"
ORIGIN
Alignment Scores:
Pred. No.: 1.47 Length: 1263
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x AX3167371 (1-1263)

```

```

Qy 1 PheLeuTyrlleuProArgAspValleu 9
Db 664 TTCTCTACTTGCCTGCGGAGTCCTG 690

RESULT 7
CQ731678 2053 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 17612 from Patent WO02068579.
DEFINITION CQ731678
ACCESSION CQ731678
VERSION CQ731678.1 GI:42308932
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
human exons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 17612 06-SEP-2002;
PB Corporation (NY) (US)
FEATURES
source Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 2.55 Length: 2053
Pred. No.: 47.00 Matches: 9
Score: 47.00 Conserv: 0
Percent Similarity: 100.0%
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-13 (1-9) x CQ731678 (1-2053)

Qy 1 PheLeuTyrlleuProArgAspValleu 9
Db 748 TTCTTTACTTGCCTGCGGAGTCCTG 774

RESULT 8
H85T40A 2053 bp RNA linear PRI 18-APR-2005
LOCUS Homo sapiens 5T4 gene for 5T4 oncofetal antigen.
DEFINITION H85T40A
ACCESSION 229083
VERSION 229083.1 GI:435654
KEYWORDS 5T4 gene; 5T4 oncofetal antigen.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2053)
AUTHORS Myers, K.A., Rahi-Saund, V., Davison, M.D., Young, J.A., Cheater, A.J.
and Stern, P.L.
TITLE Isolation of a cDNA encoding 5T4 oncofetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
JOURNAL J. Biol. Chem. 269 (12), 9319-9324 (1994)
PUBMED 8132670
REFERENCE 2 (bases 1 to 2053)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
FEATURES
source Location/Qualifiers
/organism="Homo sapiens"
/mol_type="other RNA"

```

```

/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label=N-flank
85..1347
/codon_start=1
/evidence=experimental
/product="5T4 oncofetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGSRGPAAGDRLRLALVLLGWSSSPTSSASFSFS
APFLAGAVSAQPPPLDQCPCALCESEARTVKVNRNLTEVPTDLPAYVRLFLTGNO
LAVLPAGAFARRPPLAEALNLGSLRDLDEVRAGAFELPSRLDLSHNLADLSPP
AFSGSNASVSAPSLVLELIINHVPPDERQNRSPFGMVVAALLAGRALQGLRLLELA
SNRFLYLPDVLQAQLPSLRHLDLSNNLSLVITYSPRNLTLESILHEDNALKVLNG
TLARQLQPHIRVFLDNPNWVCDCHMDMWTWKETEVVQCKRLTCAYPEKMRNVLC
LELNSADLDCDPIPLPSPQTSYVPLGIVLALIGAIFLLVLYLNRRKGIKQWMMHNRDCA
RDHMGVHYRYEINADPRLTNLSNSDV"
130..171
373..966
/product="Leucine rich repeat region"
/label=LRRs
966..1119
/product="LRR C-terminal flank"
/label=C-flank
1153..1215
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

ORIGIN
Alignment Scores: 2.55 Length: 2053
Pred. No.: 47.00 Matches: 9
Score: 47.00 Conserv: 0
Percent Similarity: 100.0%
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x H85T40A (1-2053)

Qy 1 PheLeuTyrlleuProArgAspValleu 9
Db 748 TTCTTTACTTGCCTGCGGAGTCCTG 774

RESULT 9
BD127282 2359 bp DNA linear PAT 18-SEP-2002
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD127282
ACCESSION BD127282
VERSION BD127282.1 GI:232222227
KEYWORDS JP 2002017375-A/2713.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2713 22-JAN-2002;
HELIIX RESEARCH INSTITUTE

```

```

COMMENT      OS Homo sapiens (human)
PN JP 2002017375-A/2713
PD 22-JAN-2002
PF 07-JUL-2000 JP 200253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
C12P21/02,C12P1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS
(424)..(1572).
FEATURES
source
1..2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.99 Length: 2359
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x BD127282 (1-2359)
Qy 1 PheLeuTyRLeuProArgAspValleu 9
|||||
Db 1087 TTCCTTACCTGCGCGGATGTGCTG 1113
RESULT 10
CQ782724 2359 bp DNA linear PAT 17-MAR-2004
LOCUS
DEFINITION Sequence 2864 from Patent EP1396543.
ACCESSION CQ782724
VERSION CQ782724.1 GI:45502667
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H. for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 2864 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source
1..2359
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAP05958.1"
/db_xref="GI:45502668"
/translations="MPGCGSRGPAAGGRLLARLALRVLTGPTDPAVVRNLPTGNG
APFLASVSAQPLPDQCPALCESEAAATVKVNRNLTVPTDPAVVRNLPTGNG
LAVLPAGAFARRPPLAEALNLSGSRLDVRAGAFPHLPRLQDLSHNPLADLSFP
AFSGNSASVAGPLVSLIINLHVPPEDERQNRSFEGVVAALIGRALQGLRLLELA
SNHFLYLPDRVLALQPLSLNLDLSNNLSVSLTVYSFRNLTHLSLHLEDNALKVLHNG
TLAELQGLPHIRVPLDNNPWVCHMADMTVTLKETEVQCKDRLTCAYPEKVRNRYL
LELSADLDCDPIPLPSLQTSYVFLGIVLALIGALFLLVLYLNKRGIKK"
CDS

```

```

ORIGIN
Alignment Scores:
Pred. No.: 2.99 Length: 2359
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x CQ782724 (1-2359)
Qy 1 PheLeuTyRLeuProArgAspValleu 9
|||||
Db 1087 TTCCTTACCTGCGCGGATGTGCTG 1113
RESULT 11
AK074786 2359 bp mRNA linear PRI 03-SEP-2002
LOCUS
DEFINITION Homo sapiens cDNA FLJ90305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AK074786
VERSION AK074786.1 GI:22760460
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
TITLE
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 2359)
Isogai,T. and Otsuki,T.
Direct Submission
TITLE
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp. Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1..2359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/notes="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"
ORIGIN
Alignment Scores:
Pred. No.: 2.99 Length: 2359
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

```


US-10-774-176-13 (1-9) x AK074786 (1-2359)	REFERENCE	1	Hominidae; Homo.
Qy	AUTHORS	Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.	
Db	TITLE	Primers for synthesizing full length cDNA clones and their use	
	JOURNAL	Research Association for Biotechnology (JP)	
	FEATURES	Location/Qualifiers	
	source	1. .2361	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
	CDS	426. .1688	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAF85961.1"	
		/db_xref="GI:45503670"	
		/translation="MPGCSRGPAAGDGRRLRLARLALVLLGWSSSSPTSSASSPSSS	
		APFLASAVSAQPLPDQPCALCESEARATKCVNRNLTEVPTDLPAYVYRNFLTGNG	
		LAVPLASAVFARPPPLAEALNALNSRLDEVRAGAFELHPSLRQLDLSHNPGLRLSPF	
		AFSGSNASVSAFPLVELILNIHVPPEDQRNRPQEGMVVAALLAGRALQGLRLLELA	
		SNHFLYLPDRDLAQLPSLRHLDSNNLSVLTVSFRLNTHLESLEHLEADNAKLVLRNG	
		TLAEQGLPHIRVFLDNPWCDCMADMTWLKETEYVQGRDLTCAYPEKMRNRLV	
		LELSADLDCPLPSPQTSFVIGLVIALIGALFLLVLYLNRRGIKKOMENIRDAC	
		RDHMEGYHYRYEINADPRLTNLSSNDV"	
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	2.99 Length: 2361	
	Score:	47.00 Matches: 9	
	Percent Similarity:	100.0% Conservative: 0	
	Best Local Similarity:	100.0% Mismatches: 0	
	Query Match:	100.0% Indels: 0	
	DB:	6 Gaps: 0	
	US-10-774-176-13 (1-9) x CQ782726 (1-2361)		
Qy	1 PheLeuTyrlleuProArgAspValleu 9		
Db	1089 TTCTTTACCTGCGCGGGATGTGCTG 1115		
	LOCUS	AX961916 2361 bp DNA linear PAT 14-JAN-2004	
	DEFINITION	Sequence 127 from Patent WO03104277.	
	ACCESSION	AX961916	
	VERSION	AX961916.1 GI:40881326	
	KEYWORDS		
	SOURCE	Homo sapiens (human)	
	ORGANISM	Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
		Hominidae; Homo.	
	REFERENCE	1	
	AUTHORS	Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.	
	TITLE	Stat6 activation gene	
	JOURNAL	Patent: WO 03104277-A 127 18-DEC-2003;	
	FEATURES	Location/Qualifiers	
	source	1. .2361	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
	CDS	426. .1688	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAF06467.1"	
		/db_xref="GI:40881327"	
		/translation="MPGCSRGPAAGDGRRLRLARLALVLLGWSSSSPTSSASSPSSS	
		APFLASAVSAQPLPDQPCALCESEARATKCVNRNLTEVPTDLPAYVYRNFLTGNG	
		LAVPLASAVFARPPPLAEALNALNSRLDEVRAGAFELHPSLRQLDLSHNPGLRLSPF	
		AFSGSNASVSAFPLVELILNIHVPPEDQRNRPQEGMVVAALLAGRALQGLRLLELA	
		SNHFLYLPDRDLAQLPSLRHLDSNNLSVLTVSFRLNTHLESLEHLEADNAKLVLRNG	
		TLAEQGLPHIRVFLDNPWCDCMADMTWLKETEYVQGRDLTCAYPEKMRNRLV	
		LELSADLDCPLPSPQTSFVIGLVIALIGALFLLVLYLNRRGIKKOMENIRDAC	
		RDHMEGYHYRYEINADPRLTNLSSNDV"	
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	2.99 Length: 2361	
	Score:	47.00 Matches: 9	
	Percent Similarity:	100.0% Conservative: 0	
	Best Local Similarity:	100.0% Mismatches: 0	
	Query Match:	100.0% Indels: 0	
	DB:	6 Gaps: 0	
	US-10-774-176-13 (1-9) x BD127283 (1-2361)		
Qy	1 PheLeuTyrlleuProArgAspValleu 9		
Db	1089 TTCTTTACCTGCGCGGGATGTGCTG 1115		
	LOCUS	CQ782726 2361 bp DNA linear PAT 17-MAR-2004	
	DEFINITION	Sequence 2866 from Patent EP1396543.	
	ACCESSION	CQ782726	
	VERSION	CQ782726.1 GI:45502669	
	KEYWORDS		
	SOURCE	Homo sapiens (human)	
	ORGANISM	Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Eukaryota; Eutheria; Euarchontoglires; Primates; Catarrhini;	
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
		Hominidae; Homo.	
	REFERENCE	1 (bases 1 to 2361)	
	AUTHORS	Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.	
	TITLE	Primer for synthesizing full-length cDNA and use thereof	
	JOURNAL	HELIIX RESEARCH INSTITUTE	
	COMMENT	OS Homo sapiens (human)	
		PN JP 2002017375-A/2714	
		PD 22-JAN-2002	
		PF 07-JUL-2000 JP 2000253172	
		PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO	
		PI ISHII,	
		PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI	
		SHINICHI KOJIMA,	
		PI TETSUJI OTSUKI, HISASHI KOGA	
		PC	
		C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC	
		10,	
		PC C12P21/02, C12Q1/68/ C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC	
		Primer for synthesizing full-length cDNA and use thereof FH key	
		Location/Qualifiers	
	FT CDS	(426). .(1685).	
		Location/Qualifiers	
	source	1. .2361	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
	ORIGIN		
	Alignment Scores:		
	Pred. No.:	2.99 Length: 2361	
	Score:	47.00 Matches: 9	
	Percent Similarity:	100.0% Conservative: 0	
	Best Local Similarity:	100.0% Mismatches: 0	
	Query Match:	100.0% Indels: 0	
	DB:	6 Gaps: 0	
	US-10-774-176-13 (1-9) x BD127283		

SNHFLYLPDRVLALQPSLRHLDLSNNLSVLSVTSFRNLTHLSLHEDNALKVLHNG
TLAELQGLPHIRVFLDNNPWVCDCHMADMTLTKETEVQKDLTCAYPEKQNRVL
LELNSADLDCDPLPSPLOTSYVFLGIVLALIGAFLLVLYLNKGIKKWHNIRDAC
RDHMEGYHYEINADPRNLNLSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 2.99 Length: 2361
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-13 (1-9) x AX961916 (1-2361)

Qy 1 PheLeuTyrlLeuProArgAspValLeu 9
Db 1089 TTCCTTTACCTGCGCGGGATGCTG 1115

RESULT 15

AK074790

LOCUS

DEFINITION Homo sapiens cDNA FLJ30309 fis, clone NT2RP2000903, highly similar to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.

AK074790

ACCESSION AK074790.1 GI:22760466

VERSION oligo capping; fis (full insert sequence).

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE

AUTHORS

Otsuki, T., Ota, T., Nishikawa, T., Hayashi, K., Suzuki, Y., Yamamoto, J., Wakamatsu, A., Kimura, K., Sakamoto, K., Hatano, N., Kawai, Y., Ishii, S., Saito, K., Kojima, S., Sugiyama, T., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Nagai, K., Sugano, S., and Isogai, T.
Signal Sequence and Keyword Trap in silico for Selection of Full-Length Human cDNAs Encoding Secretion or Membrane Proteins from Oligo-Capped cDNA Libraries
DNA Res. 12, 117-126 (2005)

JOURNAL

REFERENCE

AUTHORS

Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahari, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Ninomiya, K.
NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 2361)

Isogai, T. and Otsuki, T.

Direct Submission

Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,

Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan

(E-mail: genomics@hri.co.jp, Tel.81-438-52-3975, Fax.81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology; cDNA library construction:

Institute of Medical Science, University of Tokyo, Laboratory of

Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass

sequencing and clone selection: Helix Research Institute (supported

by Japan Key Technology Center etc.).

Location/Qualifiers

1. .2361

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP2000903"

/cell_line="NT2"

/cell_type="teratocarcinoma"

FEATURES

source

/clone lib="NT2RP2"
/note="Cloning vector: pME18SPL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 2.99 Length: 2361
Score: 47.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x AK074790 (1-2361)

Qy 1 PheLeuTyrlLeuProArgAspValLeu 9
Db 1089 TTCCTTTACCTGCGCGGGATGCTG 1115

RESULT 16

BC037161

LOCUS

DEFINITION Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317 IMAGE:4138906), complete cds.

BC037161

ACCESSION BC037161.2 GI:33872201

VERSION MGC.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 2379)
Strausberg, R.L., Feingold, S.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, J., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uadin, T.B., Tothiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalka, U., Smalil, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

2 (bases 1 to 2379)
Strausberg, R.
Direct Submission
Submitted (03-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URI: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:22713382.
Contact: MGC help desk
Email: cgabs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

cDNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mcgenhgr.nih.gov

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 26 Row: m Column: 15

This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5729717.

FEATURES

source

1. .2379

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:15317 IMAGE:4138906"
 /tissue_type="Muscle, rhadomyosarcoma"
 /clone_lib="NIH MGC 17"
 /lab_host="DH10B-R"
 /notes="Vector: pOTB7"

gene

1. .2379

/gene="TPBG"
 /notes="synonyms: MGP1, 574-AG, 574"
 /db_xref="GeneID:7162"
 /db_xref="MIM:190920"

CDS

427..1689

/gene="TPBG"
 /codon_start=1
 /product="574 oncofetal trophoblast glycoprotein"
 /protein_id="AAH37161.1"
 /db_xref="GI:22713383"
 /db_xref="GeneID:7162"
 /db_xref="MIM:190920"
 /translation="MPGCSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSS
 APFLASAAQPPLPDQPCALCESEARTVKCNRLNLTVEPTDLPAYVNLFLTGQ
 LAVLPAAGAFRRPLALNLSGRLEDEVRAGFEHLPSRLQDLSHNPLADLSPP
 SNFSLVSDVAPSLVLELNLHVPPEDEQRNSFGVMVAALLAGRALQGLRLSLA
 SHFLVPRVLAQLPSLRLHLSNLSLSTVVSFRLNLTSLSLHEDNALVLENG
 TLAEGLQHLRHFVLDNNPVCCHMADMTWLKEFVCGDKRLTCAYPEKMRNVL
 LEINSDLCDDPLPSPSLQTSYVFLGIVLALIGLIFLLVLYLNRKGIKKMMHNRDAC
 RDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
 Pred. No.: 3.02 Length: 2379
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x BC037161 (1-2379)

QY 1 PheLeuTyrlLeuProArgAspValLeu 9

Db 1090 TTCTTTACCTGCGGGATGCTG 1116

RESULT 17

AB168308

LOCUS

DEFINITION

AB168308

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AB168308 2714 bp mRNA linear PRI 18-JUN-2005
 Macaca fascicularis testis cDNA clone: Qtsa-11109, similar to human
 trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.

AB168308.1 GI:67967899
 oligo capping; fis (full insert sequence).
 Macaca fascicularis (crab-eating macaque)
 Macaca fascicularis

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Cercopitheidae; Cercopithecinae; Macaca.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

VECTOR: pCYPAC2

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source

1..121909
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="q13-15"
 /clone="RP3-492P14"
 /clone_lib="RPCI-3"
 100
 /notes="Clone right end: RP1-93K22"
 complement(1004..10982)
 /locus tag="RP3-492P14.2-001"
 /locus tag="RP3-492P14.2-001"
 /pseudoc
 complement(1004..10982)
 /locus tag="RP3-492P14.2-001"
 /notes="match: proteins: P81877 Q99LX9 Q9BWM6 Q9CY28 Q9D614 Q9P38 Q9Y4T7"
 /pseudoc
 /codon_start=1
 86539
 /notes="Clone left end: RP1-90G1"
 109639..116836
 /gene="TPBG"
 /locus tag="RP3-492P14.1-001"
 join(109639..109916,110631..116836)
 /gene="TPBG"
 /locus tag="RP3-492P14.1-001"
 /product="trophoblast glycoprotein"
 /notes="match: ESTs: AA149121 AA152323 AA565852 AA643734 AL544610 AW471072 AW662538 BE260089 BF306457 BF306926 BF314984 BI196133 B1562387 BM069833 BM670613 match: cDNAs: AJ420536.1 Z29083.1 110970..112232
 /gene="TPBG"
 /locus tag="RP3-492P14.1-001"
 /standard_name="OTTHUMP00000016786"
 /notes="match: proteins: Q13641 Q9QYD9 Q9Z0L0"
 /codon_start=1
 /product="trophoblast glycoprotein"
 /protein_id="CAI21546.1"
 /db_xref="GI:56203539"
 /db_xref="GOA:Q13641"
 /db_xref="InterPro:IPR000372"
 /db_xref="InterPro:IPR000483"
 /db_xref="InterPro:IPR001611"
 /db_xref="InterPro:IPR003591"
 /db_xref="UniProt/TREMBL:Q13641"
 /translation="MFGGSRGPAAGDGLRLRLARLALVLGWMSSSPTSSASFSSS APFLARRAQPPALQCCEAAATVKCVNRLTEVPTDLPAYVNLFLTGQ LAYLPAGAPRRPLAEALNLSGRLDEVRAGPEHLPSLQDLSHNPLADLSPF AFGSNASVAPSLVLELNLHIVPPEDEQNSFEGMVVAALLAGHALQGLRLLELA SNHFLYLPDRVLQALSRLHLDLSNNLSVLTIVSFNLTLSLESLDNLKVLKNG TLABLQGLPHIRVFLDNNPWCDCHMADMTWLKETEVSQGDKRLTCAYPEKMRNVL LKNSADLDCDPLDLPSPSLQTSYVFLGIVLALIGAILFLVLYLNRKGIKKWMHNRDAC RDHMEGYHYRYEINADPRLTNLSNSDV"
 116817..116822
 /gene="TPBG"

polyA_signal

/locus_tag="RP3-492P14.1-001"
 116836
 /gene="TPBG"
 /locus_tag="RP3-492P14.1-001"
 121909
 /note="Clone_right_end: RP3-492P14"
 ORIGIN
 Alignment Scores:
 Pred. No.: 264 Length: 121909
 Score: 47.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-13 (1-9) x HSJ492P14 (1-121909)
 Qy 1 PhLeuTyrieuProArgAspValLeu 9
 |||||
 Db 111633 TTCCTTTACCTGCCGCGGATGTGCTG 111659
 RESULT 20
 AF527803 28770 bp DNA linear PRI 23-OCT-2003
 LOCUS Homo sapiens cyclin-dependent kinase inhibitor 2A (CDKN2A) gene,
 DEFINITION complete cds, alternatively spliced.
 AF527803
 ACCESSION AF527803.1 GI:21886808
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1 (bases 1 to 28770)
 AUTHORS Rieder,M.J., Livingston,R.J., Daniels,M.R., Montoya,M.A.,
 Chung,M.-W., Miyamoto,K.E., Nguyen,C.P., Nguyen,D.A., Poel,C.L.,
 Robertson,P.D., Schackwitz,W.S., Sherwood,J.K., Witrak,L.A. and
 Nickerson,D.A.
 DIRECT SUBMISSION
 SUBMITTED (08-JUL-2002) Genome Sciences, University of Washington,
 1705 NE Pacific, Seattle, WA 98195, USA
 COMMENT To cite this work please use: NIEHS-SNPs, Environmental Genome
 Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
 (URL: <http://egp.gs.washington.edu>).
 FEATURES
 Location/Qualifiers
 1..28770
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 196..27063
 /gene="CDKN2A"
 join(196..673,23606..23912,26572..27063)
 /gene="CDKN2A"
 /product="cyclin-dependent kinase inhibitor 2A, p14ARF"
 join(358..673,23606..23811)
 /gene="CDKN2A"
 /notes="alternatively spliced"
 /codon_start=1
 /product="cyclin-dependent kinase inhibitor 2A, p14ARF"
 /protein_id="AAW77919.1"
 /db_xref="GI:21886809"
 /translation="MGRGRCVGPSLQRLQGRWCRCPLVPKGGAAAAELGPGGNNWVR
 RFLVTLIRACGPVRVVFVHIFRLTGEWAFGAPAAVALVLMRLSRQLGQPLP
 RRPCHDQGRPSGGAARRAQLRRPHSHPTRARRCPGGLPGHAGGAAPGRGAAG
 RARCIGPSARGPG"
 529
 /gene="CDKN2A"
 /frequency="0.01"
 /replace="t"
 1763
 /gene="CDKN2A"
 variation
 variation

misc_feature	/frequency="0.10"	/replace="t"	variation	/replace="n"
	1958..2120	/gene="CDKN2A"		7229
variation	/note="Region not scanned for variation"			/gene="CDKN2A"
	2537	/frequency="0.02"	variation	/replace="g"
	/replace="t"			7291
variation	2888	/gene="CDKN2A"		/gene="CDKN2A"
	/frequency="0.43"	/replace="a"	variation	/frequency="0.01"
	3059	/gene="CDKN2A"		/replace="g"
	/frequency="0.02"	/replace="c"		7341
variation	3116	/gene="CDKN2A"		/frequency="0.05"
	/frequency="0.01"	/replace="c"	variation	/replace="a"
	3159	/gene="CDKN2A"		7460
	/frequency="0.01"	/replace="g"	variation	/frequency="0.03"
repeat_region	3398..3616	/rpt_family="MER1_type"		/replace="t"
	/rpt_type=dispersed		variation	7658
variation	3440	/gene="CDKN2A"		/frequency="0.01"
	/frequency="0.48"	/replace="a"	variation	/replace="n"
	3512..4111	/gene="CDKN2A"		7706..7708
misc_feature	/note="Region not scanned for variation"			/gene="CDKN2A"
	3895..4107	/rpt_family="L1"	repeat_region	/frequency="0.01"
	/rpt_type=dispersed			7954
misc_feature	4188..4519	/gene="CDKN2A"		/frequency="0.01"
	/note="Region not scanned for variation"		variation	/replace="a"
variation	5334	/gene="CDKN2A"		7966
	/frequency="0.04"	/replace="g"		/gene="CDKN2A"
variation	5481	/gene="CDKN2A"		/frequency="0.23"
	/frequency="0.01"	/replace="g"	variation	/replace="a"
	5718	/gene="CDKN2A"		8153..8273
	/frequency="0.98"	/replace="n"		/rpt_family="MIR"
variation	5916	/gene="CDKN2A"		/rpt_type=dispersed
	/frequency="0.11"	/replace="g"		8542
variation	6012	/gene="CDKN2A"		/gene="CDKN2A"
	/frequency="0.03"	/replace="a"	variation	/frequency="0.01"
	6044..6254	/rpt_family="L1"		/replace="c"
repeat_region	/rpt_type=dispersed			8595
	6334..6979	/gene="CDKN2A"		/gene="CDKN2A"
misc_feature	/note="Region not scanned for variation"			/frequency="0.05"
	6819..6970	/rpt_family="MER1_type"	variation	/replace="a"
	/rpt_type=dispersed			8845
variation	7227	/gene="CDKN2A"		/frequency="0.01"
	/frequency="0.97"		variation	/replace="g"
				9042
				/gene="CDKN2A"
				/frequency="0.01"
				/replace="c"
				9237
				/gene="CDKN2A"
				/frequency="0.01"
				/replace="t"
				10152
				/gene="CDKN2A"
				/frequency="0.10"
				/replace="g"
				10482
				/gene="CDKN2A"
				/frequency="0.01"
				/replace="t"
				10727
				/gene="CDKN2A"
				/frequency="0.01"
				/replace="g"
				10756..10935
				/rpt_family="MIR"
				/rpt_type=dispersed
				10794
				/gene="CDKN2A"

```

variation
/frequency="0.02"
/replaces="a"
10803
/gene="CDKN2A"
/frequency="0.01"

Alignment Scores:
Pred. No.: 87.6 Length: 28770
Score: 46.00 Matches: 8
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 97.9% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x AF527803 (1-28770)

Qy 1 PheLeuTyrieuProArgAspValleu 9
Db 14453 TTCTGTATCTCCCGAGATATTCTA 14479

RESULT 21
AC000048/c
LOCUS
DEFINITION
AC000048 Homo sapiens Chromosome 9p21 Cosmid Clone c66, complete sequence.
ACCESSION
VERSION AC000048.4 GI:5882760
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 34669)
AUTHORS
Burian,D.M., Mitchell,N. and Roe,B.A.
TITLE
Homo sapiens Cosmid Clone c66 encoding the p16/CDK-INK4 gene
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 34669)
AUTHORS
Sveen,L., Olopade,F.I. and Rowley,J.D.
JOURNAL
Unpublished
REFERENCE
3 (bases 1 to 34669)
AUTHORS
Roe,B.A.
JOURNAL
Direct Submission
TITLE
Submitted (30-OCT-1996) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
4 (bases 1 to 34669)
AUTHORS
Roe,B.A.
JOURNAL
Direct Submission
TITLE
Submitted (15-SEP-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
5 (bases 1 to 34669)
AUTHORS
Roe,B.A.
JOURNAL
Direct Submission
TITLE
Submitted (09-APR-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Sep 15, 1999 this sequence version replaced gi:5801678.

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES
source
Location/Qualifiers
1..34669
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chr="9"
/map="9p21"
/clone="c66"
/notes="Left Neighbor sequenced in this center: 34a5. Right
Neighbor sequenced in this center: c86."

ORIGIN
Alignment Scores:
Pred. No.: 108 Length: 34669
Score: 46.00 Matches: 8
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 97.9% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x AC000048 (1-34669)

Qy 1 PheLeuTyrieuProArgAspValleu 9
Db 34279 TTCTGTATCTCCCGAGATATTCTA 34253

RESULT 22
AL449423/c
LOCUS
DEFINITION
AL449423 101155 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone RP11-14912 on chromosome 9 Contains
the CDKN2A gene for cyclin-dependent kinase inhibitor 2A (melanoma,
p16, inhibits CDK4), the gene for susceptibility protein NSG-x
(LOC51198), the 5' end of a variant of the MTAP gene for
methylthiodenosine phosphorylase (MSAP), the 3' end of the CDKN2B
gene for cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)
and 4 CpG islands, complete sequence.
ACCESSION
VERSION AL449423.14 GI:16944057
KEYWORDS
HTG; CDK4; CDKN2A; CDKN2B; CpG island; LOC511198; MSAP; MTAP; NSG-x.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 101155)
AUTHORS
Babbage,A.
JOURNAL
Direct Submission
TITLE
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegasanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Nov 15, 2001 this sequence version replaced gi:15795445.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-14912 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VBCOR: pBACE3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegasanger.ac.uk

-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES
source
Location/Qualifiers
1..101155
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"

```

misc_feature	/clone="RP11-149I2" /clone_lib="RPCI-11.1" 2000 /notes="Clone right_end: RP11-70L8" 57670..58492 /gene="C9orf53" /locus_tag="RP11-149I2.3-001" 57670..58492 /genes="C9orf53" /locus_tag="RP11-149I2.3-001" /product="chromosome 9 open reading frame 53" /notes="match: ESTs: AA729227.1 A4883806.1 AA897353.1 AA922526.1 AI806497.1 BE85998.1 match: cDNAs: AF211119.1" 57941..58180 /genes="C9orf53" /locus_tag="RP11-149I2.3-001" /standard_name="OTHUMP0000021151" /notes="match: proteins: Q9UH64" /codon_start=1 /product="chromosome 9 open reading frame 53" /protein_id="CAH70598.1" /db_xref="GI:55665305" /translations="MEQRGGEHLPTSVKSEPRACNPNPTVAENRVPSGLAARVRLTA LWNPSLVGVSERRGGDWSPRIPLWARVGWIQLPG" 58474..58479 /genes="C9orf53" /locus_tag="RP11-149I2.3-001"	CDS	/locus_tag="RP11-149I2.2-005" /product="cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)" /note="match: ESTs: BE393890.1 BF974706.1" complement(join(58970..58983,61643..61949,65419..65568)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-001" /standard_name="OTHUMP0000021147" /note="match: proteins: BAB91133 P42771 Q9NP05" /codon_start=1 /product="cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)" /protein_id="CAH70600.1" /db_xref="GI:55665307" /translation="MEPAAGSMESPADWLATAARGRVEEVRALEALGNAPNSY GRRTIOVMMSGASVARELLLLHGAEPCNDPATLTRPVHDAAREGFDTLVLHRRAGA RLDVRDAGRLPVDLAERLGHDRDVARYLRAGGTSGSNHARIDAEAGSDIPD" complement(join(60310..60534,61643..61949)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-008" complement(join(60310..60534,61643..61949)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-008" /product="cyclin-dependent Kinase inhibitor 2A (melanoma, p16, inhibits CDK4)" /note="match: ESTs: AA909181.1 AA973236.1 AI989792.1" complement(join(61643..61949,85562..86042)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-006" complement(join(61643..61949,85562..86042)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-006" /product="cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)" /note="match: ESTs: BG178804.1" complement(join(61643..61949,84880..85213)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-002" complement(join(61643..61949,84880..85213)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-002" /product="cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)" /note="match: cDNAs: S78535.1" complement(join(61744..61949,84880..85195)) /gene="CDKN2A" /locus_tag="RP11-149I2.2-002" /standard_name="OTHUMP0000021148" /codon_start=1 /product="cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)" /protein_id="CAH70601.1" /db_xref="GI:55665308" /db_xref="GOA:Q8N726" /db_xref="InterPro:IPRO10868" /translation="WGRCVCVPISQLSQRWCSCSLVPKGGAARALGPGGCVNRVR RFLVTILIRACGPRVFVVHIPIELTGWNAAPAAVAVALVLLRSQRLCQQQPIL RRPGHDGQRSFGAAAAARRGAQLRRPHSHPTARRCPGGLPGHAGGAOPRGANG RARCLGPSARGPF" complement(65218..65568) /gene="CDKN2A" /locus_tag="RP11-149I2.2-003" /standard_name="OTHUMP0000021149" /codon_start=1 /product="cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)" /protein_id="CAH70599.1" /db_xref="GI:55665306" /db_xref="GOA:Q5VVJ3" /db_xref="InterPro:IPRO02110"
gene		gene	
mRNA		mRNA	
CDS		CDS	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	
gene		gene	
mRNA		mRNA	
polyA_site		polyA_site	
polyA_signal		polyA_signal	

$\alpha\gamma$

AC092291 194551 bp DNA linear PRI 19-MAR-2003
 LOCUS Homo sapiens chromosome 16 clone CTD-2502C11, complete sequence.
 AC092291
 AC092291.3 GI:29124062
 HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE
 AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
 Alamos National Laboratory.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (29-SEP-2001) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (19-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 COMMENT On Mar 19, 2003 this sequence version replaced gi:15808524.
 Draft Sequence produced by DOE Joint Genome Institute
 www.igi.doe.gov
 Finishing Completed at Stanford Human Genome Center and Los Alamos
 National Laboratory
 www.bhgsc.stanford.edu
 Quality: Phrap Quality >=40 99.8% of Sequence;
 Estimated Total Number of Errors is 0.4.

FEATURES
 source
 1..194551
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="CTD-2502C11"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.32e+03 Length: 194551
 Score: 45.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.7% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-13 (1-9) x AC092291 (1-194551)

Qy 1 PheLeuTyrlieuProArgAspValleu 9
 |||::|||
 Db 20510 TTTATCTACCTACCCCGTGATGCTT 20536

RESULT 26
 AF063939
 LOCUS Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA, ROD 01-JAN-2000
 DEFINITION complete cds.
 ACCESSION AF063939
 VERSION AF063939.1 GI:6650211
 KEYWORDS Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
 Ninkina, N.N. and Buchman, V.L.
 Structure and expression of the rat 5T4 gene
 Unpublished
 2 (bases 1 to 2333)
 Buchman, V.L.
 Direct Submission
 Submitted (06-MAY-1998) School of Biomedical Sciences, University
 of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
 UK

FEATURES
 source

Location/Qualifiers
 1..2333
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /tissue_type="cerebellum"
 /dev_stage="newborn"
 1..2333
 /gene="5T4"
 1..363
 /gene="5T4"
 364..1644
 /gene="5T4"
 /codon_start=1
 /product="5T4 oncofetal antigen homolog"
 /protein_id="AAP21770.1"
 /db_xref="GI:6650212"
 /translation="MPGAGSGFSGAGDGRLLRLARLALLVGLWVSAPSSSLPSSSTQ
 PAAFLASGSAQPPPAERCPAACESEARTVKVNRNLLEVPADLPYYVNRNLTGNQ
 MTVLPAGAFARQPLADLAVLNLGNHLKEVGAGAFELPLGLRLDLDSNPLTNLSAF
 TPAGSNVSVSTSPILLELTLNHIVPEDQRQSGFEGMVAFGMWAAALRSGLALRGAL
 HLELASHPFLVLPDLDLPSLAKHLDRNNSLYSLTYASFNLTHLESLHLEADNAL
 KVHNSHTLAWQGLAHVRVFLDNNPVCCTNADVMWSLKEFTEVPDKKRLTCAPFEK
 MERNGLDITSDLDCCDATLPQSLOTSTVYFLGIVLALIGAIPLLVLVLYLNKRGIKKWMH
 NTRDRCRDHMEGYHYREINADPSLTNLSNSGV"
 1645..2333
 /gene="5T4"
 2315..2320
 /gene="5T4"

3'UTR
 polyA_signal

ORIGIN

Alignment Scores:
 Pred. No.: 14.8 Length: 2333
 Score: 44.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 93.6% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-13 (1-9) x AF063939 (1-2333)

Qy 1 PheLeuTyrlLeuProArgaspValleu 9
 |||||
 Db 1045 TTTCCTTACCTGCTGCTGGACTTATTG 1071

RESULT 27
 BC087011
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BC087011
 Rattus norvegicus trophoblast glycoprotein, mRNA (cDNA clone
 MGC:93332 IMAGE:7193411), complete cds.
 BC087011
 BC087011.1 GI:56268819
 MGC.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
 1 (bases 1 to 2361)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D., Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toehyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McRwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A., 2002. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 99 (26), 16899-16903 (2002) 12477932

2 (bases 1 to 2361)
Director MGC Project.
Direct Submission
Submitted (02-DEC-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 186 Row: 0 Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

Location/Qualifiers
1. .2361
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="MGC:9332 IMAGE:7193411"
/tissue_type="Heart, rat (Brown Norway)"
/clone_lib="NIH MGC_234"
/lab_host="DH10B"
/notes="Vector: pExpress1"
1. .2361
/gene="Tpbp"
/notes="synonym: 5T4"
/db_xref="GeneID:83684"
/db_xref="RGD:621453"
364..1644
/gene="Tpbp"
/codon_start=1
/product="Tpbp protein"
/protein_id="AAH87011.1"
/db_xref="GI:56268820"
/db_xref="GeneID:83684"
/db_xref="RGD:621453"
/translations="MFGAGRGSGAGGRLRLRLRLALVLGWVSAPSSSLPSSSTS
PAAFLASGAQPPAPCAACSEARTVKVCNRLLEVPADLPFYVRLNLTGQ
MTVLPAGAPQPLADLAVNLISGNHLKEVGAGAFHPLGLRLDLSHNPLTNLSAF
TFAGSVSVSTPSPLELLILNIVPPEDQKNGSFEGWAFEGWAAALRLSGLRLRL

gene

CDS

HHLELASNHFLYLPDRLLDQLPSIKHLDRNSLVSTYASFRNLTHLESILHEDNAL
KVLHNSLTAEWQGLAHVRVFLDNNPNWCCDCMADMSWLKETEVPVPOKARTCAPPEK
MRRNGLLDTSSDLDCDAPLQSLQTSYVFLGIVLALIGALFLVLYLNKRGIKQWMMH
NIRDACDRHMEGYHYRVEINADPRLTNLSSNDV"

ORIGIN

Alignment Scores:
Pred. No.: 15 Length: 2361
Score: 44.00 Matches: 8
Percent Similarity: 100.0% Conservatism: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.6% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-13 (1-9) x BC0807011 (1-2361)

Qy 1 PhleuTyrieuProArgAspValLeu 9

Db 1045 TTTCCTTACCTGCTCGGACTTATTG 1071

RESULT 28

AC128294/c
LOCUS AC128294 210237 bp DNA linear HTG 19-NOV-2002
DEFINITION Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.
ACCESSION AC128294
VERSION AC128294.3 GI:25083347
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 210237)
Muzny, D., Marle, Metzger, M., Lees, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Blawlo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, L., Louised, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C., Plummer, F., Poinexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rokey, T., Rojars, A., Rose, R., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpungsoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Popper, P., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojasa, A., Rose, M., Rose, K., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 239076)
Worley, K.C.

Direct Submission
Submitted (14-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 239076)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:22857070.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GOPI
Center clone name: CH230-87110

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 228642 bases at least Q40
Consensus quality: 232269 bases at least Q30
Consensus quality: 234041 bases at least Q20
Estimated insert size: 231522; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 234710: contig of 234710 bp in length

* 234711 234810: gap of unknown length
* 234811 235924: contig of 1114 bp in length
* 235925 236024: gap of unknown length
* 236025 237314: contig of 1290 bp in length
* 237315 237414: gap of unknown length
* 237415 239076: contig of 1662 bp in length.

FEATURES
source
1. 239076
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-87110"
234711..234810
/estimated_length=unknown
235925..236024
/estimated_length=unknown
237315..237414
/estimated_length=unknown

ORIGIN
Alignment Scores:
Pred. No.: 2.84e+03 Length: 239076
Score: 44.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.6% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-13 (1-9) x AC106962 (1-239076)
Qy 1 PhleuTyrieuProArgAspValleu 9
|||||
Db 16048 TTCTTACCTGCTCGGACTATTG 16022

RESULT 30
AX467373 1260 bp DNA linear PAT 16-JUL-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.

REFERENCE
1
AUTHORS
TITLE
JOURNAL
Myers, K., Drury, N. and Carroll, M.
Polypeptide
Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
1. 1260
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"

ORIGIN
Alignment Scores:
Pred. No.: 12.6 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-13 (1-9) x AX467373 (1-1260)
Qy 1 PhleuTyrieuProArgAspValleu 9
|||||
Db 661 TTCTTCTTCTGCTCGGACGTACTG 687

RESULT 31

```

AX821533
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS Felis catus (cat)
SOURCE
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
AUTHORS Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
TITLE MHC class I peptide epitopes from the human St4 tumor-associated
antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 12.6 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x AX821533 (1-1260)
Qy 1 PheLeuTyrlEuProArgAspValleu 9
|||||:|||||:|||||:|||||:|||||
Db 661 TTCTCTTCTGCTCGGACGTACTG 687

RESULT 32
AX821548
LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS Felis catus (cat)
SOURCE
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
AUTHORS Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
TITLE MHC class II peptide epitope of St4 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 12.6 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-13 (1-9) x AX821548 (1-1260)
Qy 1 PheLeuTyrlEuProArgAspValleu 9

```

```

|||||:|||||:|||||:|||||:|||||
Db 661 TTCTCTTCTGCTCGGACGTACTG 687

RESULT 33
U40187/c
LOCUS U40187 26599 bp DNA linear INV 21-SEP-2004
DEFINITION Caenorhabditis elegans cosmid F11H8, complete sequence.
ACCESSION U40187
VERSION U40187.4 GI:46195904
KEYWORDS HTG.
SOURCE
ORGANISM Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 26599)
.
WormBase Consortium
Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
Science 282 (5396), 2012-2018 (1998)
9851916
REFERENCE
AUTHORS Menezes,S. and Menezes,S.
CONSRTM The sequence of C. elegans cosmid F11H8
TITLE Unpublished (2001)
JOURNAL
AUTHORS
REFERENCE 3 (bases 1 to 26599)
TITLE Waterston,R.
JOURNAL
AUTHORS
REFERENCE Direct Submission
TITLE Submitted (06-NOV-1995) Robert Waterston
JOURNAL
AUTHORS
REFERENCE 4 (bases 1 to 26599)
TITLE Waterston,R.
JOURNAL
AUTHORS
REFERENCE Direct Submission
TITLE Submitted (28-JUN-2001) Department of Genetics, Washington
JOURNAL University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
AUTHORS
REFERENCE 5 (bases 1 to 26599)
TITLE Waterston,R.
JOURNAL
AUTHORS
REFERENCE Direct Submission
TITLE Submitted (19-APR-2002) Department of Genetics, Washington
JOURNAL University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
AUTHORS
REFERENCE 6 (bases 1 to 26599)
TITLE Waterston,R.
JOURNAL
AUTHORS
REFERENCE Direct Submission
TITLE Submitted (19-NOV-2002) Department of Genetics, Washington
JOURNAL University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
AUTHORS
REFERENCE 7 (bases 1 to 26599)
TITLE Wilson,R.
JOURNAL
AUTHORS
REFERENCE Direct Submission
TITLE Submitted (07-JUL-2003) Department of Genetics, Washington
JOURNAL University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
AUTHORS
REFERENCE 8 (bases 1 to 26599)
TITLE Wilson,R.
JOURNAL
AUTHORS
REFERENCE Direct Submission
TITLE Submitted (05-APR-2004) Department of Genetics, Washington
JOURNAL University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
AUTHORS
REFERENCE 9 (bases 1 to 26599)
TITLE WormBase Consortium
CONSRTM
TITLE Direct Submission
JOURNAL
AUTHORS
REFERENCE Submitted (21-SEP-2004) Department of Genetics, Washington
JOURNAL University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
COMMENT On Apr 5, 2004 this sequence version replaced gi:20198805.
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis , MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England

```

email: submissions@watson.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

For a graphical representation of this clone sequence and its analysis see:

<http://www.wormbase.org/db/seq/sequence?name=Fl1H8;class=Sequence>

NEIGHBORING CLONE INFORMATION

The 5' clone is C06E8, 2400 bp overlap; the 3' clone is R01H2, 200 bp overlap. Actual start of this clone is at base position 2397 of Fl1H8; actual end is at 11501 of R01H2.

NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program GeneFinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFeome cloning project (<http://wofdb.dfci.harvard.edu/>), similarity to other proteins from BlastX analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES

```

source
1. .26599
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="III"
/clone="Fl1H8"
35. .10194
/genes="Fl1H8.2"
/locus_tag="Fl1H8.2"
join(35..91,376..437,5100..5194,5867..5934,5986..6249,
6556..6694,6888..7091,7266..7353,7401..7514,7569..7664,
8148..8215,9879..10019,10111..10194)
/genes="Fl1H8.2"
/locus_tag="Fl1H8.2"
/standard_name="Fl1H8.2"
/notes="contains similarity to Pfam domains PF02931
(Neurotransmitter-gated ion-channel ligand binding
domain), PF02932 (Neurotransmitter-gated ion-channel
transmembrane region); coded for by the following C.
elegans cDNAs: yk1165g06.3, yk1297c02.3, yk1297c02.5,
yk1335f08.3"
/codon_start=1
/product="Hypothetical protein Fl1H8.2"
/protein_id="AA881158.3"
/db_xref="GI:46195905"
/db_xref="WormBase:Fl1H8.2"
/translation="MTWLLSLLIGVYTHPEHPPLMKSDKLLIYAPQASRTEPTLGGSEN
ENLENNVRLDGLIASHDRIRPVGPPPIVNTAHVTIISAVSEVSDYTDLLFL
RQIWDPLRAWSDSDSLTGIDMVKTIWPTDTPFNKSKSFHEATNSFLRIND
HGVLKSLRITVANCPSLHTFLDRQCALEVESYGSKDIIYHHCHTNVATIDE
NVHLAHTTGEKHCHERTSLSTGNTYSRLTAYFLKRNIGFYLIQYFPSSLIWIWS
VSWPLNREAVQARVAIGVTIVLTWTLTSTNASLTPKVSIVKSLDVLGVCFIVPAS

```

gene

complement(10409..11354)
/genes="col-8"

CDS

/locus_tag="Fl1H8.3"
complement(join(10409..10480,10530..10909,10958..11354))
/genes="col-8"
/locus_tag="Fl1H8.3"
/standard_name="Fl1H8.3"
/notes="contains similarity to Pfam domains PF01391
(Collagen triple helix repeat (20 copies)), PF01484
(Nematode cuticle collagen N-terminal domain); coded for
by the following C. elegans cDNAs: CK575722, OSTF154C2_1,
OSTR154C2_1, yk69e3.3, yk69e3.5, yk72h10.3, yk72h10.5,
yk223a2.3, yk223a2.5, yk233e6.3, yk233e6.5, yk280b6.3,
yk280b6.5, yk381e2.3, yk381e2.5, yk388a4.3, yk388a4.5,
yk724h6.3, yk724h6.5, yk738g10.3, yk738g10.5"
/codon_start=1
/product="Collagen protein 8"
/protein_id="AA881159.1"
/db_xref="GI:1055196"
/db_xref="WormBase:Fl1H8.3"
/translation="MLVCVFVLYTMGLLTDIKQLQSDDEDFEFAITKDTWQRI
VTXHTYPGVDEETIESHPPTFTLFGTRKARQAYPSQCNGPKSEGCAPGPPGGE
GGSGEPGHDGDKPGAGVIVAITHDIPGCIKCPGRPGRPGSLVGPAGPAG
QGRGPPGTGGGGGCGEAGRPGGAAGRPGPRGEGTGYRQAGRAGPPGR
GPQPEGNPGAGEDGNQGVGHGVPGRPGIPGKSGTCGHHGPGRPGDAGVCPCP
GRSYKA"
complement(11391..11392)
/notes="8L1 trans-splice acceptor; see yk724h6.5"
12674..12675
/notes="8L1 trans-splice acceptor; see yk705c7.5"
12680..14356
/genes="rf1-1"
/locus_tag="Fl1H8.1"
join(12680..12862,12908..12990,13040..13120,13190..13298,
13347..13608,13696..13897,13941..14117,14161..14356)
/genes="rf1-1"
/locus_tag="Fl1H8.1"
/standard_name="Fl1H8.1"
/notes="contains similarity to Pfam domains PF00899 (ThiP
family), PF02134 (Repeat in ubiquitin-activating (UBA)
protein); coded for by the following C. elegans cDNAs:
CK587479, OSTF034Fl1_1, OSTR034Fl1_1, yk705c7.3,
yk705c7.5"
/codon_start=1
/product="Ectopic membrane ruffles in embryo protein 1"
/protein_id="AA881160.2"
/db_xref="GI:32453024"
/db_xref="WormBase:Fl1H8.1"
/translation="MVSVDPLATERWRSIRLTDSDAYKVPWFVQGPENFEALQNTK
ILVIGAGLCEKLNALSGFRTEIVYDMTDIVSNLRQFLFREDVGSKAEVAA
AFVQRVVGCQVTAHNCRIEDKGQFYKFSIIICGLDSDPAREWINGMLCDVLEMA
DGKPDENTIIIMIDGGTGEGKNARVYPKFTACIDCTLDLPQVNPFLCTIATPR
LPHCIRYIKVVMVPEKPFEGVSLDADDPHVEWLVLERASLRAEKNIKRGVDRRLTS
GVLKRITPAVASTNAVIAASCALEALKLTATNIAKPIDNINLFTOIHGAYTSVSNMKD
DNCLTCGGRLPFEVSPSSLESILRLSERFLNKHPTLATSTRKLYCISSEFMPQFEQ
ESKENLHSMKDLVSDGGEILVSDALSRLTLRIQLI"

Alignment Scores:

Pred. No.:	402	Length:	26599
Score:	43.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	91.5%	Indels:	0
DB:	2	Gaps:	0

US-10-774-176-13 (1-9) x U40187 (1-26599)

Oy 1 PheLeuTyrLeuProArgAspValLeu 9

Db 7888 TTTTGTACCTGCCACGACATAGTA 7862

RESULT 34	AY275838/c	27297 bp	DNA	linear	BCT 24-FEB-2004
LOCUS	Escherichia coli strain Cl3 serovar O113:H21 pathogenicity island				
DEFINITION	I, partial sequence.				
ACCESSION	AY275838				
VERSION	AY275838.1	GI:33440434			
KEYWORDS	Escherichia coli				
SOURCE	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.				
ORGANISM	1 (bases 1 to 27297)				
REFERENCE	Shen,S., Macarennas,M., Rahn,K., Kaper,J.B. and Karmali,M.A.				
AUTHORS	Evidence for a Hybrid Genomic Island in Verocytotoxin-Producing				
TITLE	Escherichia coli Cl3 (Serotype O113:H21) Containing Segments of				
JOURNAL	BDL933 (Serotype O157:H7) O Islands 122 and 48				
PUBMED	Infect. Immun. 72 (3), 1496-1503 (2004)				
REFERENCE	14977955				
AUTHORS	2 (bases 1 to 27297)				
TITLE	Shen,S., Macarennas,M., Rahn,K., Kaper,J. and Karmali,M.A.				
JOURNAL	Direct Submission				
FEATURES	Submitted (14-APR-2003) Laboratory for Foodborne Zoonoses, Health				
source	Canada, 110 Stone Road West, Guelph, Ontario N1G 3W4, Canada				
	Location/Qualifiers				
	1. .27297				
	/organism="Escherichia coli"				
	/mol_type="genomic DNA"				
	/strain="Cl3"				
	/serovar="O113:H21"				
	/db_xref="taxon:562"				
	28. .651				
	/genes="Z1635"				
	28. .651				
	/genes="Z1635"				
	/note="Similar to Escherichia coli O157:H7 EDL933 Z1635				
	gene encoded by GenBank Accession Number AE005311"				
	/codon_start=1				
	/transl_table=11				
	/product="Z1635-like protein"				
	/protein_id="AAQ19121.1"				
	/db_xref="GI:33440435"				
	/translation="MTRIDYADKYCTERNESPCILQQQMDVLECRQTEAGPEER				
	LRLLANVDYVTFELPRLTRTPOLIAALREWDISQNVVFNDRKFCVYSLSKA				
	SLSGVPTFHYLSHRFRRVVGNETSLPYQOAVREKAPRERLKVALEAGLLVTLTD				
	GLWFQSORAADVLRKAGMPVTTTVEVDNLTCTTRKVPAYHL"				
	747. .980				
	/genes="Z1636"				
	747. .980				
	/genes="Z1636"				
	/note="Similar to Escherichia coli O157:H7 EDL933 Z1636				
	gene encoded by GenBank Accession Number AE005311"				
	/codon_start=1				
	/transl_table=11				
	/product="Z1636-like protein"				
	/protein_id="AAQ19122.1"				
	/db_xref="GI:33440436"				
	/translation="MARSXDYVNLSEHLEHNEWLAIGVAGLVDRNRLAEVVRKL				
	QDSFYINVDALNTAYSEHPFWSGLVSGDEN"				
	complement (1033. .1224)				
	/genes="Z1637"				
	complement (1033. .1224)				
	/genes="Z1637"				
	/note="Similar to Escherichia coli O157:H7 EDL933 Z1637				
	gene encoded by GenBank Accession Number AE005311"				
	/codon_start=1				
	/transl_table=11				
	/product="Z1637-like protein"				
	/protein_id="AAQ19123.1"				
	/db_xref="GI:33440437"				
	/translation="MSEDVPLPKVQRYKODHGLAVTVTSVEETRVVFMREGYPHPCM				
	RPWYNPLGKPKPEBTE"				
gene	1900. .3552				
CDS	/gene="S1"				
	1900. .3552				
	/gene="S1"				
	/note="Similar to Escherichia coli O157:H7 EDL933 Z1640				
	gene encoded by GenBank Accession Number AE005312 and to				
	Yersinia pestis CO92 putative hemolysin activator protein				
	encoded by GenBank Accession Number CAC91296"				
	/codon_start=1				
	/transl_table=11				
	/product="putative hemolysin activator HlyB"				
	/protein_id="AAQ19124.1"				
	/db_xref="GI:33440438"				
	/translation="MSVLLSPLSLQAADIRBSGEDTFIIQQORQALBQOLTPSPADV				
	RLSPGFAHKINFVETPCFIQIKTELKGADALPHWLPLQKIANGAVGCHCLGAKGIN				
	LLMSTQLRLVDHGVTVTRVLAPSDQLKSGILRLVIFGVVNRHLTPDSDYIQIYS				
	SFPAHEGSLDLRDIQGLENNQRLPGVQADMEIMPQPGESDILITRKQDKYRWLG				
	PSLDGSTRSGRYLGGITFSLDNPFSLSLLYVSATNFPHYGGKSGKNYTAHVSVP				
	FGYWFSTASDIDYHOTVAGVEDYQYSGESQNLGMOLSRVLHRNGTKTTLTVDVQ				
	ARRSRYNIDREIQVRQTAARLGLQHRHYISQATLDAGVSWQGRFQWFGQAPAE				
	EIFGRATALSILQNNASFNDPLTAEQPFSTFRIYQLSNTPTPQBPQALGNWT				
	VRGDFERTLNASNGWYIRNDLAWRTPLPQELYLGMQYGEVGGAGSEYLVGRHLAGG				
	VVGLRGQLFSTDYDVFIGRPFSPKAGFMTSDVTAGNFLFRDP"				
	3562. .4089				
	/gene="S2"				
	3562. .4089				
	/gene="S2"				
	/note="Similar to Ralstonia solanacearum putative				
	hemolysin activating-like protein encoded by GenBank				
	Accession Number CAD18225"				
	/codon_start=1				
	/transl_table=11				
	/product="putative hemolysin activator HlyC"				
	/protein_id="AAQ19125.1"				
	/db_xref="GI:33440439"				
	/translation="MRNGKYDVPLYSGEVNEAEVLGAAVLAWHSPHLDAPLHT				
	LPDLLPVIRKRYVATEQGRPVFFMSQAWLSPAEARYLTQAILPQSDWNSGDR				
	MWCWDVAPFGHTFDMGRNLVRQDIFPRLCWRSLYHRGQQTGLRVNMFQAGVSRBEAR				
	QMRVQHLPAVDVPER"				
	4156. .6006				
	/gene="S3"				
	4156. .6006				
	/gene="S3"				
	/note="Similar to Yersinia pestis putative hemolysin and				
	putative adhesin encoded by GenBank Accession Number				
	AJ414152 and to Yersinia pestis CO92 putative adhesin				
	encoded by GenBank Accession Number AL590842"				
	/codon_start=1				
	/transl_table=11				
	/product="putative hemolysin/hemagglutinin"				
	/protein_id="AAQ19126.1"				
	/db_xref="GI:33440440"				
	/translation="MVVADIARSGAGTSLSRGCGYRQRICRVTPLAFSLWLAGSM				
	VHSVNAAGIAADHAGAPHQQTITQTASGIFQVNIQTSAGVSHNTYSQPDVGNKV				
	ILNNAHNVQTLGCVAGNPLAKGEARIILNEVNSPQLNGFVEVAGKAQVVI				
	ANPAGISCDGCGFINARATLTITQPMKNGSLTGFVERGEIQTITGKMDASRTDYT				
	DI IARSVKINAGIWAQDLKVTTRNNVDIAHQTEKKAADASSQVQALDVSLGMY				
	AGKRLTGTETGVGRNAGHIGAQAGAVTLTADGRIENSGSI SAKTDVHLATTRLHN				
	SGSVYACQDTIQSNGVPTHTGVSANRNTRIQTARLTGGERSLLAAGVDDGRLLAA				
	GNLVTSTTGELAAHGQVLGQDLGSLNSRIQOQHTELDATSGNLSTQNVQL				
	SAGTLSARTAGHNGNGGTINADTLQISAQSLNHRGLIQGTGDFSLNLPGGVNR				
	EGLLAANGAVRLDALSLNDRGKQVAVQSGSLQVKTGTAVDNQCGSLTASRDVRLNAQ				
	ALNNDNGLISAAGTGRIKTQQAVSNTGEMESAGRLGTSAGSLNNHQV"				
	6027. .13058				
	/gene="S4"				
	6027. .13058				
	/gene="S4"				
	/note="Similar to Yersinia pestis CO92 putative hemolysin				
	and putative adhesin encoded by GenBank Accession Number				
	AJ414152 and to Yersinia pestis CO92 putative adhesin				
	encoded by GenBank Accession Number AL590842"				
	/codon_start=1				
	/transl_table=11				


```
/product="putative adhesin/hemagglutinin/hemolysin"  
/protein_id="AA019127.1"  
/db_xref="GI:3340441"  
/translation="MSDGLSVTLGDALDNTSGRLLSQKTLSSVSGSELVSDGLTQSGS  
DMLTDVQDGLVLRNRTKRGISAVPYSAGMLNQCQFIVGQKMTLNAAGTLTQSGS  
GVLSQASLIQISGTLNQKALKAGTDMLLSGDVSNGEGTLAAGEDLNAHLNVLEN  
QQTIVSNGSRVDVTFDNDQGRVVAQSLTSSDIIINDASGLTQSGASLNRADT  
LGNRSGDGRVLSQSPMTLNAGTLDSTAGVLSGALSITAGVNNSTQGVANGLL  
GWNQALNQSGLIQGRG19INTAGQTLNRRGTLNSQLBELTSTGMDNRGVTGAK  
TTADLTSLDNREGRLVSEGLRLHTGQSHGQIQSGVMDLNSRVGVNDVSG  
LIPSGA1LTNALQFTNRHTQNTQGLIACQTHITTDLDNORGSILADRALVTMADR  
TLNNPQVLSGATLVSQRQLAFNRDGVVKAGQSVSDVDAQGLGDKGLSLGMDTL  
KLTPSNGSGTTLANGNLTLVSGDVSNTSLGSLRDLNLSIRLENTEGESISAGQT  
WLVNTDNLGRGLIDKQTRLQNTLNTSGTRIYGDVAGVSAGATFNHLEENGATYTL  
AGRVDVLTQTLNNRTHSLIYAGDMHTGMLDANGAATGKAGVLNHNSEATYAGY  
LVLQAQNNVNDHFTTRVVTVEKTEVQLSGDKRWSAGRPVGVVNDNSNSLKK  
LHTEGARDKFTQDTRTVEETRVESDPKILSGAGMTIIVADKLVNDNSQVAGGL  
LTIIPSGSVSVSGRHTVDSGTSTYYIRKKGDKQKQSGTSQYTPPVITQITLTK  
PGLTSHGVQVQSHVTLSPKQGTVDVQTLGNVDATVAGTDRIPLRPVVSAGEPVI  
LLPQOQFVSAPQGSIHVAGPDLRLPDSLLFTKNPAVNPVLYETDPFRFNQTKWLS  
DYMKAFSONGMLKRLGDGFYQRLIREQVVALTGQRLDGVSNDEQFKALMDAG  
IARQKYNLTGVALTAQOMALTGDIVLWNTVTLPGSTOTVQVPOVYARVKPD  
VNSAGAL1AGRDVWKLDGLFNSGKLAKQTQLSARNHNQAGTIQGANVSLTART  
DINSTGLGLQATDLSLAMAGRD1SLFTTTRTAQRDAGQNHFTSIDSVAGVYQNDQ  
GRVLQAGRDMLNTAAATVNVQKDSLTLQAGRDMLTSTVTSADQNTWIDKNNRSLDE  
GVTQSTGSLAGMDVTLTAGRDWTSQAASLSAQKGLALMAGHDVTLTGAQNTSRLDE  
YHKVTGSGSLKNTTTHDVTDRRTTSGELNGDVTISAGHNLNATGSSVAGDNRV  
SLAAGNNL1GTLTENRETHLQKSKSLGSLMSSGGVQSPSGSKVTDATDTTQKG  
STVSGVSDVLSQAGNRLTVNSDLTAGRDMLSKGKSVISAATDQVQTHTEQKTS  
GUTLALSGTAGSLNLTVEVQAASAGNSRLALQVKAALSGAQAORAVOAGLADAG  
ADAGNNNTV1SLSYQSGSQQSBEQTVAGKSTLTAGNNLSITQATSGVGVGDGLSV  
TIGS01 KAGNN1LQANRDVNLVSAKNTSKLGGKNTSGGSGVGVGVGVGVGLSV  
BASANGKSGKNGTHTEVTDAGNRLTIIISGRDPTLGAQAGETVKVDAGRLT  
LTSQSDR1DSKQNASAGSTFFGSMGSAVNLKDKHNSYDSVQQTGIFAGR  
GGSDVTGQHTLQNGVIASTADTKNRLDTGLTGFSDIENRADFKTEHQSAGLSTGG  
SVAGNTLGMANNL1GANHEHADSTTQSAVAGNITIRDTQSKQDVADLKRDAH  
ANOTLSP1PREKBEQQAOLIGIETGNQVADIARTEGQIAGEKAKRDPAALNOARA  
ELEAKGPTTEODVQARVYNNMAAPRGVGGYQQAIOAATAVQGLAGNLSAALA  
GGAVP1AE1IKQTPDAGRVATHAVVNAALSLAQKNALAGAGNAGTEVVGMLAT  
QMTGKPSVSE1ETKQGLVATVAGLGLVGDGSAVAGQSGKTTVENNALSF  
GDGPENAAATSWNKYA VDNNTTPQQTQAGLDKIAKQMDPDSTNITKVIYDQYQGV  
MTAGVYLAASAGKVLGGGLGLAANSQYQIYDLNQPNQANKSWDYLGSATSFYTG  
MMAPDRGLVANTG1AMGGAFTDQPNPTASLAGAGIAGLGAAGKVAFTAVGVLGND  
PVPQF1VELGGVASEFTNGI1KDFNPPQAPKKNK"  
13446..13670  
/gene="85"  
13446..13670  
/gene="85"  
/notes="similar to Versinia pestis CO92 putative adhesin  
encoded by GenBank Accession Number AL590842"  
/codon_start=1  
/transl_table=11  
/product="putative adhesin-like factor"  
/protein_id="AA019128.1"  
/db_xref="GI:3340442"  
/translation="MVNENQLGSGSEWLQTEKAREHGAYVLSGSDNPSGACKRQGA  
HKYAAALATGGCYQGVPKL1CGYLVPEQP"
```

Alignment Scores:

Pred. No.:	414	Length:	27297
Score:	43.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	91.5%	Indels:	0
DB:	1	Gaps:	0

US-10-774-176-13 (1-9) x AY275938 (1-27297)

Qy 1 PheLeuTyrLeuProArgAspValLeu 9

Db 23637 TTCTTTATCTGCTAAGGATATTC 23611

RESULT 35

```
BD184766/c  
LOCUS  
DEFINITION  
Nucleic acid molecule and polypeptide specific to intestinal  
hemorrhagic pathogenic Escherichia coli O157:H7, and method of use  
thereof.  
ACCESSION  
BD184766  
VERSION  
BD184766.1  
KEYWORDS  
GI:31876966  
SOURCE  
JP 2002355074-A/2  
ORGANISM  
Escherichia coli O157:H7  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
REFERENCE  
1 (bases 1 to 86248)  
AUTHORS  
Hayashi,H., Shinagawa,H., Makino,K., Hayashi,T., Onishi,S.,  
Hattori,M. and Kurokawa,K.  
TITLE  
Nucleic acid molecule and polypeptide specific to intestinal  
hemorrhagic pathogenic Escherichia coli O157:H7, and method of use  
thereof.  
JOURNAL  
Patent: JP 2002355074-A 2 10-DEC-2002;  
PRESIDENT OF UNIVERSITY OF TSUKUBA  
COMMENT  
OS Escherichia coli O157:H7  
PN JP 2002355074-A/2  
PD 10-DEC-2002  
PF 24-JAN-2002 JP 2002015959  
PI HIDEO HAYASHI,HIDRO SHINAGAWA,KOZO MAKINO,TETSUYA HAYASHI,SHIN  
ONISHI,  
MASAHIRA HATTORI,KEN KUROKAWA  
PC C12N15/09,C12N15/09,A61K31/7088,A61K39/00,A61K48/00,A61P31/04,  
C07K14/245,  
PC C07K16/12,C12M1/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/  
02,C12Q1/68,  
PC G01N33/15,G01N33/50,G01N33/53,G01N33/56,G01N37/00,  
C12N15/00,  
PC C12N15/00,C12N5/00  
CC Nucleic acid molecule and polypeptide specific to intestinal  
hemorrhagic  
CC pathogenic Escherichia coli O157:H7, and method of use thereof  
FH Key Location/Qualifiers  
FT source  
1..86248  
/organism="Escherichia coli O157:H7".  
Location/Qualifiers  
1..86248  
/organism="Escherichia coli O157:H7"  
/mol_type="genomic DNA"  
/db_xref="taxon:83334"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1..53e+03 Length: 86248  
Score: 43.00 Matches: 7  
Percent Similarity: 100.0% Conservative: 2  
Best Local Similarity: 77.8% Mismatches: 0  
Query Match: 91.5% Indels: 0  
DB: 6 Gaps: 0  
US-10-774-176-13 (1-9) x BD184766 (1-86248)  
Qy 1 PheLeuTyrLeuProArgAspValLeu 9  
Db 64602 TTCTTTATCTGCTAAGGATATTC 64576  
RESULT 36  
AR204161/c  
LOCUS  
DEFINITION  
Sequence 57 from patent US 6365723.  
ACCESSION  
AR204161  
VERSION  
AR204161.1  
KEYWORDS  
GI:21500738  
ORGANISM  
Unknown.  
REFERENCE  
1 (bases 1 to 87563)  
AUTHORS  
Blattner,P.R., Burland,V., Perna,N.T., Plunkett,G. and Welch,R.
```

TITLE Sequences of E. coli O157
JOURNAL Patent: US 6365723-A 57 02-APR-2002;
FEATURES Location/Qualifiers
source
1. .87563
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores: Length: 87563
Pred. No.: 1.56e+03
Score: 43.00
Percent Similarity: 100.0%
Best Local Similarity: 77.8%
Query Match: 91.5%
DB: 6

US-10-774-176-13 (1-9) x AR204161 (1-87563)

QY 1 PheLeuTyrlleuProArgAspValleu 9
|||||TTCCTTATCGCTAGGATATTC 64564

Db 64590 TTCCTTATCGCTAGGATATTC 64564

RESULT 37

AR637510/c AR637510 87563 bp DNA linear PAT 20-APR-2005

LOCUS AR637510 Sequence 57 from patent US 6855814.

DEFINITION AR637510

VERSION AR637510.1 GI:62771252

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 87563)

AUTHORS Blattner, F.R., Burland, V., Perna, N.T., Plunkett, G. and Welch, R.

TITLE Sequences of E. coli O157

JOURNAL Patent: US 6855814-A 57 15-FEB-2005;

Wisconsin Alumni Research Foundation; Madison, WI

FEATURES

source
1. .87563
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores: Length: 87563
Pred. No.: 1.56e+03
Score: 43.00
Percent Similarity: 100.0%
Best Local Similarity: 77.8%
Query Match: 91.5%
DB: 6

US-10-774-176-13 (1-9) x AR637510 (1-87563)

QY 1 PheLeuTyrlleuProArgAspValleu 9
|||||TTCCTTATCGCTAGGATATTC 64564

Db 64590 TTCCTTATCGCTAGGATATTC 64564

RESULT 38

AE005174_11/c

WPCOMMENT Sequence split into 56 fragments LOCUS AE005174 Accession AE005174

Fragment Name Begin End
AE005174_00 1 110000
AE005174_01 100001 210000
AE005174_02 200001 310000
AE005174_03 300001 410000
AE005174_04 400001 510000
AE005174_05 500001 610000
AE005174_06 600001 710000
AE005174_07 700001 810000
AE005174_08 800001 910000
AE005174_09 900001 1010000
AE005174_10 1000001 1110000

Fragment Name Begin End
AE005174_00 1 110000
AE005174_01 100001 210000
AE005174_02 200001 310000
AE005174_03 300001 410000
AE005174_04 400001 510000
AE005174_05 500001 610000
AE005174_06 600001 710000

AE005174_11 1100001 1210000
AE005174_12 1200001 1310000
AE005174_13 1300001 1410000
AE005174_14 1400001 1510000
AE005174_15 1500001 1610000
AE005174_16 1600001 1710000
AE005174_17 1700001 1810000
AE005174_18 1800001 1910000
AE005174_19 1900001 2010000
AE005174_20 2000001 2110000
AE005174_21 2100001 2210000
AE005174_22 2200001 2310000
AE005174_23 2300001 2410000
AE005174_24 2400001 2510000
AE005174_25 2500001 2610000
AE005174_26 2600001 2710000
AE005174_27 2700001 2810000
AE005174_28 2800001 2910000
AE005174_29 2900001 3010000
AE005174_30 3000001 3110000
AE005174_31 3100001 3210000
AE005174_32 3200001 3310000
AE005174_33 3300001 3410000
AE005174_34 3400001 3510000
AE005174_35 3500001 3610000
AE005174_36 3600001 3710000
AE005174_37 3700001 3810000
AE005174_38 3800001 3910000
AE005174_39 3900001 4010000
AE005174_40 4000001 4110000
AE005174_41 4100001 4210000
AE005174_42 4200001 4310000
AE005174_43 4300001 4410000
AE005174_44 4400001 4510000
AE005174_45 4500001 4610000
AE005174_46 4600001 4710000
AE005174_47 4700001 4810000
AE005174_48 4800001 4910000
AE005174_49 4900001 5010000
AE005174_50 5000001 5110000
AE005174_51 5100001 5210000
AE005174_52 5200001 5310000
AE005174_53 5300001 5410000
AE005174_54 5400001 5510000
AE005174_55 5500001 5528445

Continuation (12 of 56) of AE005174 from base 1100001 (AE005174 Escherichia coli O157:H7)

Alignment Scores: Length: 110000
Pred. No.: 2.02e+03
Score: 43.00
Percent Similarity: 100.0%
Best Local Similarity: 77.8%
Query Match: 91.5%
DB: 1

US-10-774-176-13 (1-9) x AE005174_11 (1-110000)

QY 1 PheLeuTyrlleuProArgAspValleu 9

Db 23224 TTCCTTATCGCTAGGATATTC 23198

RESULT 39

AE005174_15/c

WPCOMMENT

Sequence split into 56 fragments LOCUS AE005174 Accession AE005174

Fragment Name Begin End

AE005174_00 1 110000
AE005174_01 100001 210000
AE005174_02 200001 310000
AE005174_03 300001 410000
AE005174_04 400001 510000
AE005174_05 500001 610000
AE005174_06 600001 710000

AE005174_07 700001 810000
AE005174_08 800001 910000
AE005174_09 900001 1010000
AE005174_10 1000001 1110000
AE005174_11 1100001 1210000
AE005174_12 1200001 1310000
AE005174_13 1300001 1410000
AE005174_14 1400001 1510000
AE005174_15 1500001 1610000
AE005174_16 1600001 1710000
AE005174_17 1700001 1810000
AE005174_18 1800001 1910000
AE005174_19 1900001 2010000
AE005174_20 2000001 2110000
AE005174_21 2100001 2210000
AE005174_22 2200001 2310000
AE005174_23 2300001 2410000
AE005174_24 2400001 2510000
AE005174_25 2500001 2610000
AE005174_26 2600001 2710000
AE005174_27 2700001 2810000
AE005174_28 2800001 2910000
AE005174_29 2900001 3010000
AE005174_30 3000001 3110000
AE005174_31 3100001 3210000
AE005174_32 3200001 3310000
AE005174_33 3300001 3410000
AE005174_34 3400001 3510000
AE005174_35 3500001 3610000
AE005174_36 3600001 3710000
AE005174_37 3700001 3810000
AE005174_38 3800001 3910000
AE005174_39 3900001 4010000
AE005174_40 4000001 4110000
AE005174_41 4100001 4210000
AE005174_42 4200001 4310000
AE005174_43 4300001 4410000
AE005174_44 4400001 4510000
AE005174_45 4500001 4610000
AE005174_46 4600001 4710000
AE005174_47 4700001 4810000
AE005174_48 4800001 4910000
AE005174_49 4900001 5010000
AE005174_50 5000001 5110000
AE005174_51 5100001 5210000
AE005174_52 5200001 5310000
AE005174_53 5300001 5410000
AE005174_54 5400001 5510000
AE005174_55 5500001 5528445

Continuation (16 of 56) of AE005174 from base 1500001 (AE005174 Escherichia coli O157:H7)

Alignment Scores:
Pred. No.: 2.02e+03 Length: 110000
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 1 Gaps: 0

US-10-774-176-13 (1-9) x AE005174_15 (1-110000)

Qy 1 PheLeuTyrLeuProArgAspValleu 9
Db 18831 TTTCCTTATCTGCCTAAGGATATTC 18805

RESULT 40
BA000007_14/c
WPCOMMENT

Sequence split into 55 fragments LOCUS BA000007 Accession BA000007
Fragment Name Begin End
BA000007_00 1 110000
BA000007_01 100001 210000
BA000007_02 200001 310000

BA0000007_03 300001 410000
BA0000007_04 400001 510000
BA0000007_05 500001 610000
BA0000007_06 600001 710000
BA0000007_07 700001 810000
BA0000007_08 800001 910000
BA0000007_09 900001 1010000
BA0000007_10 1000001 1110000
BA0000007_11 1100001 1210000
BA0000007_12 1200001 1310000
BA0000007_13 1300001 1410000
BA0000007_14 1400001 1510000
BA0000007_15 1500001 1610000
BA0000007_16 1600001 1710000
BA0000007_17 1700001 1810000
BA0000007_18 1800001 1910000
BA0000007_19 1900001 2010000
BA0000007_20 2000001 2110000
BA0000007_21 2100001 2210000
BA0000007_22 2200001 2310000
BA0000007_23 2300001 2410000
BA0000007_24 2400001 2510000
BA0000007_25 2500001 2610000
BA0000007_26 2600001 2710000
BA0000007_27 2700001 2810000
BA0000007_28 2800001 2910000
BA0000007_29 2900001 3010000
BA0000007_30 3000001 3110000
BA0000007_31 3100001 3210000
BA0000007_32 3200001 3310000
BA0000007_33 3300001 3410000
BA0000007_34 3400001 3510000
BA0000007_35 3500001 3610000
BA0000007_36 3600001 3710000
BA0000007_37 3700001 3810000
BA0000007_38 3800001 3910000
BA0000007_39 3900001 4010000
BA0000007_40 4000001 4110000
BA0000007_41 4100001 4210000
BA0000007_42 4200001 4310000
BA0000007_43 4300001 4410000
BA0000007_44 4400001 4510000
BA0000007_45 4500001 4610000
BA0000007_46 4600001 4710000
BA0000007_47 4700001 4810000
BA0000007_48 4800001 4910000
BA0000007_49 4900001 5010000
BA0000007_50 5000001 5110000
BA0000007_51 5100001 5210000
BA0000007_52 5200001 5310000
BA0000007_53 5300001 5410000
BA0000007_54 5400001 5498450

Continuation (15 of 55) of BA000007 from base 1400001 (BA000007 Escherichia coli O157:H7)

Alignment Scores:
Pred. No.: 2.02e+03 Length: 110000
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 1 Gaps: 0

US-10-774-176-13 (1-9) x BA000007_14 (1-110000)

Qy 1 PheLeuTyrLeuProArgAspValleu 9
Db 35058 TTTCCTTATCTGCCTAAGGATATTC 35032

RESULT 41
AC091229_04
WPCOMMENT

Sequence split into 12 fragments LOCUS AC091229 Accession AC091229
Fragment Name Begin End

```

AC091229_00      1      110000
AC091229_01      100001      210000
AC091229_02      200001      310000
AC091229_03      300001      410000
AC091229_04      400001      510000
AC091229_05      500001      610000
AC091229_06      600001      710000
AC091229_07      700001      810000
AC091229_08      800001      910000
AC091229_09      900001     1010000
AC091229_10     1000001     1110000
AC091229_11     1100001     1188748
Continuation (5 of 12) of AC091229 from base 400001 (AC091229 Rattus norvegicus clone CH2)

Alignment Scores:
Pred. No.:      2.02e+03      Length:      110000
Score:          43.00      Matches:      8
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    91.5%      Indels:      0
DB:             14      Gaps:        0

US-10-774-176-13 (1-9) x AC091229_04 (1-110000)

Qy      1 PheLeuTyRleuProArgAspVal 8
Db      11763 TTCTGTATCTTCCTAGGGATGTG 11786

RESULT 42
WPCOMMENT
Sequence split into 7 fragments      LOCUS AC091242 Accession AC091242
Fragment Name      Begin      End
AC091242_0          1      110000
AC091242_1        100001      210000
AC091242_2        200001      310000
AC091242_3        300001      410000
AC091242_4        400001      510000
AC091242_5        500001      610000
AC091242_6        600001      671619
Continuation (5 of 7) of AC091242 from base 400001 (AC091242 Rattus norvegicus clone CH2)

Alignment Scores:
Pred. No.:      2.02e+03      Length:      110000
Score:          43.00      Matches:      8
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    91.5%      Indels:      0
DB:             14      Gaps:        0

US-10-774-176-13 (1-9) x AC091242_4 (1-110000)

Qy      1 PheLeuTyRleuProArgAspVal 8
Db      53493 TTCTGTATCTTCCTAGGGATGTG 53516

RESULT 43
WPCOMMENT
Sequence split into 6 fragments      LOCUS AC091347 Accession AC091347
Fragment Name      Begin      End
AC091347_0          1      110000
AC091347_1        100001      210000
AC091347_2        200001      310000
AC091347_3        300001      410000
AC091347_4        400001      510000
AC091347_5        500001      595391
Continuation (5 of 6) of AC091347 from base 400001 (AC091347 Rattus norvegicus clone CH2)

Alignment Scores:
Pred. No.:      2.02e+03      Length:      110000
Score:          43.00      Matches:      8
Percent Similarity: 100.0%      Conservative: 0

```

```

Best Local Similarity: 100.0%      Mismatches:      0
Query Match:          91.5%      Indels:          0
DB:                   14      Gaps:            0

US-10-774-176-13 (1-9) x AC091347_4 (1-110000)

Qy      1 PheLeuTyRleuProArgAspVal 8
Db      47503 TTCTGTATCTTCCTAGGGATGTG 47480

RESULT 44
WPCOMMENT
LOCUS      AC025359/c
DEFINITION Homo sapiens chromosome 13 clone RP11-354D13 map 13, WORKING DRAFT
ACCESSION AC025359
VERSION   AC025359.3 GI:7656790
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 134580)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens chromosome 13, clone RP11-354D13
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 134580)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgaiter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lechoczky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McSwan,P., Mihov,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mohr,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced gi:7342149.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7802
Center clone name: 354_D_13
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 121112 bases at least Q40
Consensus quality: 127983 bases at least Q30
Consensus quality: 130794 bases at least Q20
Insert size: 147000; agarose-fp
Insert size: 132480; sum-of-contigs

```

Quality coverage: 3.4 in Q20 bases; agarose-fp
Quality coverage: 3.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 22 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1069: contig of 1069 bp in length
* 1070 1169: gap of 100 bp
* 1170 2333: contig of 1164 bp in length
* 2334 2433: gap of 100 bp
* 2434 3924: contig of 1491 bp in length
* 3925 4024: gap of 100 bp
* 4025 5207: contig of 1183 bp in length
* 5208 5307: gap of 100 bp
* 5308 7841: contig of 2534 bp in length
* 7842 7941: gap of 100 bp
* 7942 9438: contig of 1497 bp in length
* 9439 9538: gap of 100 bp
* 9539 13191: contig of 3653 bp in length
* 13192 13291: gap of 100 bp
* 13292 18133: contig of 4842 bp in length
* 18134 18233: gap of 100 bp
* 18234 22320: contig of 4087 bp in length
* 22321 22420: gap of 100 bp
* 22421 27367: contig of 4947 bp in length
* 27368 27467: gap of 100 bp
* 27468 30653: contig of 3186 bp in length
* 30654 30753: gap of 100 bp
* 30754 34208: contig of 3455 bp in length
* 34209 34309: gap of 100 bp
* 34310 41351: contig of 7042 bp in length
* 41352 46299: contig of 4849 bp in length
* 46300 52614: contig of 6215 bp in length
* 52615 59921: contig of 7207 bp in length
* 59922 60021: gap of 100 bp
* 60022 65020: contig of 4998 bp in length
* 65021 74957: contig of 9837 bp in length
* 74958 75057: gap of 100 bp
* 75058 86958: contig of 11902 bp in length
* 86959 87059: gap of 100 bp
* 87060 97235: contig of 10177 bp in length
* 97236 112962: contig of 15627 bp in length
* 112963 113062: gap of 100 bp
* 113063 134580: contig of 21518 bp in length.

```

FEATURES

```

source
1. .134580
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="13"
/clone="RP11-354D13"
/clone_lib="RPC1-11 Human Male BAC"
misc_feature
1. .1069
/notes="assembly_fragment"
gap
1070. .1169
/estimated_length=100
misc_feature
1170. .2333
/notes="assembly_fragment"
gap
2334. .2433
/estimated_length=100
misc_feature
2434. .3924
/notes="assembly_fragment"

```

```

gap
3925. .4024
/estimated_length=100
misc_feature
4025. .5207
/notes="assembly_fragment"
gap
5208. .5307
/estimated_length=100
misc_feature
5308. .7841
/notes="assembly_fragment"
gap
7842. .7941
/estimated_length=100
misc_feature
7942. .9438
/notes="assembly_fragment"
gap
9439. .9538
/estimated_length=100
misc_feature
9539. .13191
/notes="assembly_fragment"
gap
13192. .13291
/estimated_length=100
misc_feature
13292. .18133
/notes="assembly_fragment"
gap
18134. .18233
/estimated_length=100
misc_feature
18234. .22320
/notes="assembly_fragment"
gap
22321. .22420
/estimated_length=100
misc_feature
22421. .27367
/notes="assembly_fragment"
gap
27368. .27467
/estimated_length=100
misc_feature
27468. .30653
/notes="assembly_fragment"
gap
30654. .30753
/estimated_length=100
misc_feature
30754. .34208
/notes="assembly_fragment"
gap
34209. .34308
/estimated_length=100
misc_feature
34309. .41350
/notes="assembly_fragment"
vector_side:left
41351. .41450
/estimated_length=100
misc_feature
41451. .46299
/notes="assembly_fragment"
gap
46300. .46399
/estimated_length=100
misc_feature
46400. .52614
/notes="assembly_fragment"
gap
52615. .52714
/estimated_length=100
misc_feature
52715. .59921
/notes="assembly_fragment"
gap
59922. .60021
/estimated_length=100
misc_feature
60022. .65019
/notes="assembly_fragment"
gap
65020. .65119
/estimated_length=100
misc_feature
65120. .74956
/notes="assembly_fragment"
gap
74957. .75056
/estimated_length=100
misc_feature
75057. .86958
/notes="assembly_fragment"

```

```

Alignment Scores:
Pred. No.: 2.53e+03 Length: 134580
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.5% Indels: 0

```

```

DB:          14          Gaps:          0
US-10-774-176-13 (1-9) x AC025359 (1-134580)
QY  1 PheLeuTyrLeuProArgAspVal 8
Db  120120 TTCTCTACCTACCCAGAGATGC 120097

RESULT 45
AC124199
LOCUS      AC124199          157393 bp      DNA      linear      ROD 08-NOV-2003
DEFINITION Mus musculus BAC clone RP23-312B17 from 8, complete sequence.
ACCESSION  AC124199
VERSION    AC124199.3  GI:25046693
KEYWORDS  HTG.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
           Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 157393)
           Cordes, M., Haglund, K., Spalding, L., Schatzkamer, K. and
           Mangiapanello, L.
           The sequence of Mus musculus BAC clone RP23-312B17
           Unpublished (2001)
REFERENCE  2 (bases 1 to 157393)
           Wilson, R.
           Sequencing of Mus musculus
           Unpublished (2001)
REFERENCE  3 (bases 1 to 157393)
           McPherson, J.D. and Waterston, R.H.
           Direct Submission
           Submitted (12-JUN-2002) Genome Sequencing Center, 4444 Forest Park
           Parkway, St. Louis, MO 63108, USA
REFERENCE  4 (bases 1 to 157393)
           McPherson, J.D. and Waterston, R.H.
           Direct Submission
           Submitted (13-SEP-2002) Genome Sequencing Center, 4444 Forest Park
           Parkway, St. Louis, MO 63108, USA
REFERENCE  5 (bases 1 to 157393)
           McPherson, J.D. and Waterston, R.H.
           Direct Submission
           Submitted (16-NOV-2002) Genome Sequencing Center, 4444 Forest Park
           Parkway, St. Louis, MO 63108, USA
REFERENCE  6 (bases 1 to 157393)
           Wilson, R.
           Direct Submission
           Submitted (08-NOV-2003) Department of Genetics, Washington
           University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
           On Nov 16, 2002 this sequence version replaced gi:22830488.
           ----- Genome Center
           Center: Washington University Genome Sequencing Center
           Center code: WUGSC
           Web site: http://genome.wustl.edu
           Contact: submissions@watson.wustl.edu
           ----- Summary Statistics
           -----
           Center project name: M_BA0312B17
           -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Osegaawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.choxi.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.choxi.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES	source
	1. 157393
	Location/Qualifiers
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="8"
	/map="8"
	/clone="RP23-312B17"
	/clone_lib="RPCI-23"
repeat_region	1295..1479
	/rpt_family="B2"
repeat_region	1465..1511
	/rpt_family="U6"
repeat_region	1512..1724
	/rpt_family="L1"
repeat_region	3820..3911
	/rpt_family="MIR"
repeat_region	3927..4137
	/rpt_family="B2"
repeat_region	4602..4859
	/rpt_family="B4"
repeat_region	7813..7918
	/rpt_family="Alu"
repeat_region	8223..8408
	/rpt_family="B2"
repeat_region	8875..8962
	/rpt_family="ERVK"
repeat_region	10567..10615
	/rpt_family="CR1"
repeat_region	11190..11511
	/rpt_family="B4"
repeat_region	11711..11814
	/rpt_family="Alu"
repeat_region	12079..12181
	/rpt_family="MIR"
repeat_region	12982..13329
	/rpt_family="RMER15"
repeat_region	16470..16752
	/rpt_family="B4"
repeat_region	18141..18293
	/rpt_family="B4"
repeat_region	18494..18561
	/rpt_family="ID"
repeat_region	18562..18951
	/rpt_family="MaLR"
repeat_region	19252..19654
	/rpt_family="MaLR"
repeat_region	19694..19928
	/rpt_family="B2"
repeat_region	19962..20105
	/rpt_family="B4"
repeat_region	23501..23670
	/rpt_family="B2"
repeat_region	23507..23578
trna	
	/product="tRNA-Ser"
	/note="Likely pseudogene (HMM Sc=33.19 / Sec struct Sc=-10.52)"
repeat_region	24108..24261

```

repeat_region /rpt_family="MER2_type"
24790..24907
/rpt_family="Alu"
repeat_region 25473..25520
/rpt_family="ERV1"
repeat_region 26007..26094
/rpt_family="MIR"
repeat_region 26762..26829
/rpt_family="ID"
repeat_region 26903..26959
/rpt_family="Alu"
repeat_region 27290..27396
/rpt_family="MER1_type"
repeat_region 29484..29658
/rpt_family="B4"
repeat_region 29563..29692
/rpt_family="Alu"
repeat_region 29646..29703
/rpt_family="B4"
repeat_region 29724..29914
/rpt_family="B2"
repeat_region 29922..30104
/rpt_family="B2"
complement(30028..30100)
/product="tRNA-Ser"
/notes="Likely pseudogene (HMM Sc=36.07 / Sec struct
Sc=7.41)"
repeat_region 30115..30319
/rpt_family="B4"
repeat_region 31496..31641
/rpt_family="B4"
repeat_region 31962..32112
/rpt_family="Alu"
repeat_region 33184..33380
/rpt_family="B2"
repeat_region 34100..34271
/rpt_family="B4"
repeat_region 34238..34291
/rpt_family="tRNA-Ala-GCY_"
repeat_region 34995..35105
/rpt_family="MIR"
repeat_region 35106..35184
/rpt_family="Alu"
repeat_region 36306..36366
/rpt_family="T2_type"
repeat_region 36398..36454
/rpt_family="B4"
repeat_region 37070..37328
/rpt_family="T2_type"
repeat_region 37333..37412
/rpt_family="B4"
repeat_region 37588..37693
/rpt_family="B4"
repeat_region 37835..37909
/rpt_family="ID"
repeat_region 37996..38029
/rpt_family="B4"
repeat_region 38030..38428
/rpt_family="ERV1"
repeat_region 38429..38489
/rpt_family="B4"
repeat_region 38748..38886
/rpt_family="Alu"
repeat_region 38888..39036

```

```

Alignment Scores:
Pred. No.: 3.03e+03 Length: 157393
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: 9 Gaps: 0

```

```

US-10-774-176-13 (1-9) x AC124199 (1-157393)
Qy 1 PheLeuTyrlauProArgAspVal 8
Db 91227 TTTCTGTACCTGCTAGAGATGTG 91250

RESULT 46
AL672264/c 184940 bp DNA linear HTG 19-JUN-2002
LOCUS Mus musculus chromosome 16 clone RP23-407N15, 6 unordered pieces.
DEFINITION
ACCESSION AL672264
VERSION AL672264.2 GI:19309835
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 184940)
AUTHORS Tromans,A.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 8, 2002 this sequence version replaced gi:19089459.
COMMENT
----- Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: http://mrcseq.har.mrc.ac.uk
Contact: mouse@har.mrc.ac.uk
----- Project Information
Center project name: BM407N15
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 183774 bases at least Q40
Consensus quality: 183976 bases at least Q30
Consensus quality: 184092 bases at least Q20
Insert size: 184440; sum-of-contigs
Insert size: 175303; 5.1% error; agarose-fp
Quality coverage: 17.17x in Q20 bases; sum-of-contigs Quality
coverage: 18.06x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 41746: contig of 41746 bp in length
* 41747 41846: gap of 100 bp
* 41847 111308: contig of 69462 bp in length
* 111309 111408: gap of 100 bp
* 111409 114632: contig of 3224 bp in length
* 114633 114732: gap of 100 bp
* 114733 171805: contig of 57073 bp in length
* 171806 171905: gap of 100 bp
* 171906 175266: contig of 3361 bp in length
* 175267 175366: gap of 100 bp
* 175367 184940: contig of 9574 bp in length.
FEATURES
Location/Qualifiers
1..184940
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="16"
/clone="RP23-407N15"
/clone_lib="RPCI-23"
1..41746
/note="assembly_fragment:06656
misc_feature

```

```

fragment_chain:1"
41847. .111308
/Note="assembly_fragment:01652
fragment_chain:1"
111409. .114632
/Note="assembly_fragment:03321
fragment_chain:1"
114733. .171805
/Note="assembly_fragment:00150
fragment_chain:1"
clone_end:T7
vector_side:right"
171906. .175266
/Note="assembly_fragment:01351
fragment_chain:1"
175367. .184940
/Note="assembly_fragment:05594
fragment_chain:1"

misc_feature          3.64e+03      Length:      184940
                        43.00      Matches:      8
                        100.0%      Conservative: 0
                        100.0%      Mismatches: 0
                        91.5%      Indels:      0
                        14          Gaps:      0

Alignment Scores:
Pred. No.:          3.98e+03      Length:      200087
Score:              43.00      Matches:      8
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:        91.5%      Indels:      0
DB:                  8          Gaps:      0

US-10-774-176-13 (1-9) x AL672264 (1-184940)

Qy 1 PheLeuTyRLeuProArgAspVal 8
Db 2429 TTCTCTACTTACCAGAGATGTC 2406

RESULT 47
AL354821/c
LOCUS AL354821 200087 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP11-538N17 on chromosome 13,
complete sequence.
ACCESSION AL354821
VERSION AL354821.21 GI:14160928
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 200087)
          Lovell, J.
          Direct Submission
          Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
          Clone requests: clonerequests@sanger.ac.uk
          On May 20, 2001 this sequence version replaced gi:13396481.
          The following abbreviations are used to associate primary accession
          numbers given in the feature table with their source databases:
          En: EMBL; SW: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
          on the WORMPEP database can be found at
          http://www.sanger.ac.uk/Projects/C_elegans/wormpep
          This sequence
          was generated from part of bacterial clone contigs of human
          chromosome 13, constructed by the Sanger Centre Chromosome 13
          Mapping Group. Further information can be found at
          http://www.sanger.ac.uk/HGP/Chr13
          RP11-538N17 is from the library RPCI-11.2 constructed by the group
          of Pieter de Jong. For further details see
          http://www.chori.org/bacpac/home.htm
          VECTOR: pBACe3.6
          ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: vegas@sanger.ac.uk
          -----

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES
            Location/Qualifiers
             1..200087
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="13"
            /clone="RP11-538N17"
            /clone_lib="RPCI-11.2"
             1
            /Note="Clone_left_end: RP11-538N17"
            /Note="8302"
            /Note="Sequence from AC025359. Sequenced by WIBR."
            87190..87416
            /Note="Sequence confirmed by AC025359. Sequenced by WIBR."
            87417..87635
            /Note="Sequence from AC025359. Sequenced by WIBR."
            200087
            /Note="Clone_right_end: RP11-538N17"

ORIGIN
Alignment Scores:
Pred. No.:          3.98e+03      Length:      200087
Score:              43.00      Matches:      8
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:        91.5%      Indels:      0
DB:                  8          Gaps:      0

US-10-774-176-13 (1-9) x AL354821 (1-200087)

Qy 1 PheLeuTyRLeuProArgAspVal 8
Db 122398 TTCTCTACTTACCAGAGATGTC 122375

RESULT 48
AL354821/c
LOCUS AL354821 200087 bp DNA linear HTG 11-OCT-2002
DEFINITION Rattus norvegicus clone CH230-249B3, WORKING DRAFT SEQUENCE.
ACCESSION AL354821
VERSION AL354821.4 GI:23603872
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 200087)
          Muzny, D., Marie, E., Metzker, M., Lee, A., Adams, C., Alder, J.,
          Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
          Anyalabechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
          Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F.,
          Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
          Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
          Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
          Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
          Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
          Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
          Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
          Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,
          Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
          Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
          Fraser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
          Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
          Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
          Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
          Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogues, M.,

```


Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowalski, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhsuwa, L., Loulseg, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Milosavljevic, A., Miner, G., Minja, B., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervils, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpungsoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Direct Submission

Unpublished
2 (bases 1 to 210180)

Worley, K.C.

Direct Submission

Submitted (13-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 210180)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Oct. 9, 2002 this sequence version replaced gi:21741386.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GRNK

Center clone name: CH230-248B3

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 196785 bases at least Q40

Consensus quality: 198352 bases at least Q30

Consensus quality: 199387 bases at least Q20

Estimated insert size: 200819; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 210180: contig of 210180 bp in length.

FEATURES

source

1. 210180

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clones="CH230-248B3"

1. 1230

/note="wgs end extension"

clone end:Sp6"

complement(3405..4052)

/note="clone boundary"

clone end:Sp6"

site:MboI

end sequence:RXABG02TV"

177175..177338

/note="clone boundary"

clone end:T7

site:MboI

end sequence:RXABG02TJ"

ORIGIN

Alignment Scores:

Pred. No.: 4.21e+03 Length: 210180
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.5% Indels: 0
DB: Gaps: 0

US-10-774-176-13 (1-9) x AC110466 (1-210180)

Qy 1 PhleuTyriLeuProArgAspVal 8

|||||TTCCTGTATCTTCTCTAGGAGTGTG 200565

Db 200542 TTCCTGTATCTTCTCTAGGAGTGTG 200565

RESULT 49

AL672244/c

LOCUS

DEFINITION

Mouse DNA sequence from clone RP23-4318 on chromosome 16, complete

sequence.

AL672244

ACCSSION

AL672244.15 GI:28193396

VERSION

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 243114)

AUTHORS

Kay, M.

TITLE

Direct Submission

JOURNAL

Submitted (04-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquy@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk

On Feb 1, 2003 this sequence version replaced gi:24527417.

Sequence from the Mouse Genome Sequencing Consortium whole genome

shotgun may have been used to confirm this sequence. Sequence data

from the whole genome shotgun alone has only been used where it has

a phred quality of at least 30.

----- Genome Center
Center: UK Medical Research Council

Center code: UK-MRC
 Web site: <http://mrcseq.har.mrc.ac.uk>
 Contact: mouseq@har.mrc.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-4318 is from the RPI-23 Mouse BAC library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES

source

1. Location/Qualifiers

1. 243114
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosomes="16"
 /clone="RP23-4318"
 /clone_lib="RPI-23"

ORIGIN

Alignment Scores:
 Pred. No.: 4.96e+03 Length: 243114
 Score: 43.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.5% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-13 (1-9) x AL672244 (1-243114)

Qy 1 PhLeuTyLeuProArgApVal 8

Db 72633 TTCTCTACTTACCAAGAGATGTC 72610

RESULT 50

AC137320/c

LOCUS AC137320 247964 bp DNA linear HTG 20-NOV-2002
 DEFINITION Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
 *** 7 unordered pieces.

AC137320

VERSION AC137320.1 GI:25138402

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 247964)

Muzny, D., Marie, E., Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandes, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havtak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Loulseg, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 247964)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KZSP

Center clone name: CH230-unknown

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 204240 bases at least Q40

Consensus quality: 208170 bases at least Q30

Consensus quality: 211194 bases at least Q20
 Estimated insert size: 204976; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

* 1 28079: contig of 28079 bp in length
 * 28080: gap of unknown length
 * 28179: contig of 23757 bp in length
 * 51936: gap of unknown length
 * 51937: contig of 177866 bp in length
 * 229902: gap of unknown length
 * 229903: contig of 5775 bp in length
 * 230003: gap of unknown length
 * 235778: contig of 1094 bp in length
 * 235878: gap of unknown length
 * 236972: contig of 3025 bp in length
 * 237072: gap of unknown length
 * 240097: contig of 7768 bp in length.
 * 240197: 247964: contig of 7768 bp in length.

FEATURES
 source
 1. .247964
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-unknown"
 26529..28079
 /note="wgs contig"
 28080..28179
 /estimated_length=unknown
 51937..52036
 /estimated_length=unknown
 52037..53282
 /note="wgs contig"
 108194..109757
 /note="wgs contig"
 229903..230002
 /estimated_length=unknown
 230003..231101
 /note="wgs contig"
 235778..235877
 /estimated_length=unknown
 236972..237071
 /estimated_length=unknown
 240097..240196
 /estimated_length=unknown

ORIGIN

Alignment Scores:

Pred. No.:	5.07e+03	Length:	247964
Score:	43.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	91.5%	Indels:	0
DB:	14	Gaps:	0

US-10-774-176-13 (1-9) x AC137320 (1-247964)

QY 1 PheLeuTyrlieuProArgaspVal 8
 |||||
 DB 212028 TTCTGTATCTTCTAGGATGTG 212005

Search completed: April 25, 2006, 20:27:22
 Job time : 3131.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds
(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-12

Perfect score: 46

Sequence: 1 NPLTGNQL 9

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10774176/runat.24042006.165112.19185/app.query.fasta_1
-DB=N Geneseq -QFWT=fastap -SUPPIX=p2n.rng -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR_SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEA_SIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs05p
-USER=US10774176 @CGN 1.1.3463 @runat.24042006.165112.19185 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	Geneseqn1980s.*
2:	Geneseqn1990s.*
3:	Geneseqn2000s.*
4:	Geneseqn2001as.*
5:	Geneseqn2001bs.*
6:	Geneseqn2002as.*
7:	Geneseqn2002bs.*
8:	Geneseqn2003as.*
9:	Geneseqn2003bs.*
10:	Geneseqn2003cs.*
11:	Geneseqn2003ds.*
12:	Geneseqn2004as.*
13:	Geneseqn2004bs.*
14:	Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	46	100.0	343 8	ABX40605
2	46	100.0	505 14	ACL56146 Human col
3	46	100.0	927 6	ABT07721 Breast ca
4	46	100.0	927 8	ABX76333 Lung can

5	46	100.0	927	10	ADB80503	
6	46	100.0	927	11	ADN38723	
7	46	100.0	1156	6	ABV99349	
8	46	100.0	1260	6	ABK87175	
9	46	100.0	1260	10	ADB97513	
10	46	100.0	1260	10	ADB97452	
11	46	100.0	1263	3	AAA27058	
12	46	100.0	1263	4	AAF89736	
13	46	100.0	1263	6	ABK87174	
14	46	100.0	1331	8	AAD56199	
15	46	100.0	2020	10	ADJ56299	
16	46	100.0	2053	8	ACC51052	
17	46	100.0	2053	8	ABX76332	
18	46	100.0	2053	8	ABD56197	
19	46	100.0	2053	8	AAD56200	
20	46	100.0	2053	11	ADN38721	
21	46	100.0	2053	12	ADL06473	
22	46	100.0	2053	12	ADN03961	
23	46	100.0	2053	13	ADR25444	
24	46	100.0	2053	13	ACN38510	
25	46	100.0	2053	13	ADV35098	
26	46	100.0	2338	5	AA87175	
27	46	100.0	2359	4	AAK94253	
28	46	100.0	2359	12	ADL30831	
29	46	100.0	2361	4	AAK94254	
30	46	100.0	2361	12	ADL26162	
31	46	100.0	2361	12	ADL30833	
32	44	95.7	1281	3	AAA27059	
33	44	95.7	2557	12	ADL26160	
34	44	95.7	2557	12	ADL36158	
35	44	95.7	2557	12	ADO35939	
c	36	42	91.3	2820	4	ABL27914
c	37	42	91.3	5278	4	ABL27914
38	41	89.1	52754	9	ADA02798	
39	41	89.1	52754	10	ADB72536	
40	41	89.1	52754	10	ADC85278	
41	41	89.1	52754	12	ADM74393	
42	41	89.1	65274	14	ADL212653	
43	40	87.0	37091	4	ABL14244	
44	40	87.0	110000	10	ADF77343_11	
c	45	39	84.8	792	6	ABQ42606
46	39	84.8	792	6	ABQ42607	
c	47	39	84.8	795	6	ABQ42584
48	39	84.8	795	6	ABQ42585	
49	38	82.6	1874	13	ADS48405	
50	38	82.6	11646	4	AAK70925	
51	38	82.6	14148	4	AAK70926	
52	38	82.6	79590	11	ADL27152	
53	38	82.6	79684	9	ADA03074	
54	38	82.6	79684	9	ADA66358	
55	38	82.6	79684	10	ADB72812	
56	38	82.6	80988	14	ADZ12741	
57	37	80.4	240	10	ACA55519	
58	37	80.4	240	12	ADL55315	
59	37	80.4	278	10	ABX83960	
60	37	80.4	352	4	AAK57798	
61	37	80.4	390	6	ABN76160	
c	62	37	80.4	443	13	ACF82249
63	37	80.4	760	6	AA817960	
64	37	80.4	1059	3	AAA16661	
65	37	80.4	1227	8	ACA20705	
66	37	80.4	1368	8	ACA51301	
c	67	37	80.4	1368	13	ADS62019
c	68	37	80.4	1377	13	ADL11587
69	37	80.4	1526	13	ADL14741	
70	37	80.4	1573	13	ADL31327	
71	37	80.4	1614	13	ADL11499	
72	37	80.4	1680	6	ABZ12982	
73	37	80.4	1686	2	AAZ29716	
74	37	80.4	1746	10	AAZ25534	
75	37	80.4	1797	10	AAZ57635	
c	76	37	80.4	1828	12	ADQ63678
77	37	80.4	1845	13	ADL14584	

ADB80503	Ovarian c
ADN38723	Cancer/an
ABV99349	Human NOV
ABK87175	CDNA enco
ADB97513	Feline 5T
ADB97452	DNA enco
AAA27058	Human 5T4
AAF89736	Nucleotid
ABK87174	CDNA enco
ADJ56299	Human LRR
ACC51052	Human CDN
ABX76332	Lung can
ABD56197	Human LRR
AD56200	Human LRR
ADN38721	Cancer/an
ADL06473	Human tum
ADN03961	Antipsori
ADR25444	Breast ca
ACN38510	Tumour-as
ADV35098	Human CDN
AA87175	DNA enco
AAK94253	Human ful
AAK94254	Human ful
ADL26162	Human CDN
ADL30833	Full leng
AAA27059	Mouse 5T4
ADL26160	Human CDN
ADL36158	Human CDN
ADO35939	Novel mou
ABL27915	Drosophill
ABL27914	Drosophill
ADA02798	Human TNF
ADB72536	Human TNF
ADC85278	Human Tnf
ADM74393	Human car
ADL2653	Human can
ABL14244	Drosophill
Continuation (12 o	
ABQ42606	Oligonuel
ABQ42607	Oligonuel
ABQ42584	Oligonuel
ABQ42585	Oligonuel
ADS48405	Bacterial
AAK70925	Human imm
AAK70926	Human imm
ADL27152	Human gen
ADA03074	Human hCG
ADA66358	Human hCG
ADB72812	Human hCG
ADZ12741	Human can
ACA55519	Mouse sig
ADL55315	Human pol
ABX83960	Corn ear-
AAK57798	Human imm
ABN76160	Human ORP
ACF82249	Human SIR
AA817960	P. patens
AAA16661	Human sec
ACA20705	Prokaryot
ACA51301	Prokaryot
ADS62019	Bacterial
ADL11587	Plant ful
ADL14741	Plant ful
ADL31327	Plant ful
ADL1499	Plant ful
ABZ12982	Arabidops
AAZ29716	Phosphati
AAZ25534	Brassic
AAZ57635	Rice dise
ADQ63678	Novel hum
ADL14584	Plant ful

78	37	80.4	1944	8	ABX08292	Abx08292 Moss lipi	151	35	76.1	1088	14	ADY65705	Ady65705 S. manson
79	37	80.4	1962	8	ABX08291	Abx08291 Moss lipi	c 152	35	76.1	1243	6	ABL01445	AbL01445 Murine ap
80	37	80.4	2030	13	ADX03045	Adx03045 Plant ful	153	35	76.1	1326	11	ACH97244	Ach97244 Klebsiell
81	37	80.4	2139	6	AA517966	Aa517966 P. patens	154	35	76.1	1593	4	AAF63790	Aaf63790 Human sec
82	37	80.4	2172	13	ADX11901	Adx11901 Plant ful	155	35	76.1	1707	9	ADB10219	Adb10219 Alloiococ
83	37	80.4	2420	11	ADM03657	Adm03657 Human cdn	156	35	76.1	1707	9	ADB10223	Adb10223 Alloiococ
84	37	80.4	2700	11	ACL26208	AcL26208 Rice abio	157	35	76.1	1707	9	ADB10221	Adb10221 Alloiococ
85	37	80.4	3009	10	ADG32697	Adg32697 Human dna	c 158	35	76.1	1875	8	ACA24207	AcA24207 Prokaryot
86	37	80.4	3009	13	ADR14464	Adr14464 Human NF-	c 159	35	76.1	2229	11	ACH96996	Ach96996 Klebsiell
87	37	80.4	3010	10	ADP81993	Adp81993 Leukaemia	c 160	35	76.1	2235	8	ACA36080	AcA36080 Prokaryot
88	37	80.4	3418	5	ABV23380	Abv23380 Human pro	c 161	35	76.1	2671	12	ADQ35663	AdQ35663 Novel mou
89	37	80.4	3418	5	ABV23235	Abv23235 Human pro	c 162	35	76.1	2750	4	ABL23300	AbL23300 Drosophil
90	37	80.4	81748	11	ACN44090	Acn44090 Human gen	163	35	76.1	3164	4	ABL25304	AbL25304 Drosophil
91	36	78.3	240	8	ABX44575	Abx44575 Bovine ES	164	35	76.1	3229	6	ABA03002	AbA03002 Human zin
92	36	78.3	303	6	ABZ32426	Abz32426 Candida a	165	35	76.1	3313	4	ABL03253	AbL03253 Drosophil
93	36	78.3	463	8	ACA15332	AcA15332 Prokaryot	166	35	76.1	3992	4	AAH14540	AaH14540 Human cdn
94	36	78.3	463	8	ACA13740	AcA13740 Prokaryot	c 167	35	76.1	4131	10	ABZ41666	AbZ41666 N. gonorr
95	36	78.3	472	5	ABAL2530	Abal2530 Human ner	c 168	35	76.1	4173	10	ABZ41184	AbZ41184 N. gonorr
96	36	78.3	479	8	ABZ18104	Abz18104 Group III	c 169	35	76.1	4176	8	ACA41578	AcA41578 Prokaryot
97	36	78.3	506	12	ACH75203	Ach75203 Human gen	170	35	76.1	4480	14	ADY16985	Ady16985 DNA encod
98	36	78.3	597	13	ADS95969	Ads95969 Bacterial	c 171	35	76.1	4532	10	ACC49350	Acc49350 Human NET
99	36	78.3	724	3	AAW79546	AaW79546 Pinus rad	c 172	35	76.1	5047	11	ADN95651	Adn95651 Human BEC
100	36	78.3	741	4	AAH01442	AaH01442 Pseudomon	173	35	76.1	5720	4	ABL03252	AbL03252 Drosophil
101	36	78.3	1263	13	ADU25607	Adu25607 cDNA enco	174	35	76.1	6124	12	ADQ63203	AdQ63203 Novel hum
102	36	78.3	1329	8	ACA18804	AcA18804 Prokaryot	175	35	76.1	8208	4	AAK84302	AaK84302 Human imm
103	36	78.3	1386	4	AAA91315	AaA91315 NSI prote	176	35	76.1	8209	4	AAK84300	AaK84300 Human imm
104	36	78.3	1386	6	AAD36286	AdA36286 Goose par	177	35	76.1	8210	4	AAK84301	AaK84301 Human imm
105	36	78.3	1386	6	AAD46143	AdA46143 Goose par	c 178	35	76.1	15732	6	AA545389	Aa545389 Chemical
106	36	78.3	1386	6	AAD4605	AdA4605 Goose par	c 179	35	76.1	15732	6	AAK28234	AaK28234 DNA trans
107	36	78.3	1386	6	ACC69250	Acc69250 Goose par	c 180	35	76.1	15832	4	AA545399	Aa545399 Chemical
108	36	78.3	1386	10	AD140291	Ad140291 Goose par	c 181	35	76.1	15832	6	ABL33343	AbL33343 Human imm
109	36	78.3	1386	10	ABX96674	Abx96674 NSI DNA s	c 182	35	76.1	15832	6	ABK28244	AbK28244 DNA trans
110	36	78.3	1386	10	ABX96529	Abx96529 DNA encod	c 183	35	76.1	17782	3	AAA81530	AaA81530 N. mening
111	36	78.3	1884	4	AAA91313	AaA91313 Rep prote	c 184	35	76.1	27612	6	AB578852	Ab578852 E. coli C
112	36	78.3	1884	6	AAD36284	AdA36284 Goose par	c 185	35	76.1	27612	10	ADH80419	AdH80419 Escherich
113	36	78.3	1884	6	AAD46141	AdA46141 Goose par	186	35	76.1	31208	6	ABK52899	AbK52899 Genomic D
114	36	78.3	1884	6	AAD4603	AdA4603 Goose par	187	35	76.1	35133	4	ABL50991	AbL50991 Thermus c
115	36	78.3	1884	6	ACC69248	Acc69248 Goose par	188	35	76.1	35134	4	ABL50990	AbL50990 Thermus c
116	36	78.3	1884	10	AD140287	Ad140287 Goose par	c 189	35	76.1	43602	6	AB578851	Ab578851 E. coli C
117	36	78.3	1884	10	ABX96672	Abx96672 Rep DNA s	c 190	35	76.1	43602	10	ADH80418	AdH80418 Escherich
118	36	78.3	1884	10	ABX96527	Abx96527 DNA encod	c 191	35	76.1	51157	13	ABD32986	AbD32986 Mouse can
119	36	78.3	2000	11	ACL35344	AcL35344 Rice stre	c 192	35	76.1	51259	2	AAK83007	AaK83007 Partial m
120	36	78.3	2273	13	AD511585	Ad511585 Human the	c 193	35	76.1	59341	11	ACN44706	Acn44706 Human gen
121	36	78.3	2976	11	ACL26858	AcL26858 Rice abio	194	35	76.1	90541	6	AB552847	Ab552847 Human SR
122	36	78.3	3180	11	ACL28854	AcL28854 Rice abio	195	35	76.1	90541	10	ADJ37690	Adj37690 Human kin
123	36	78.3	3180	12	AD145342	Ad145342 Rice isop	196	35	76.1	90541	13	ADR31219	Adr31219 Human SRP
124	36	78.3	3483	12	ADQ25846	Adq25846 Human G-p	c 197	35	76.1	110000	3	AAA81490_01	Continuation (2 of
125	36	78.3	3517	8	AAH41182	AaH41182 Human G p	c 198	35	76.1	110000	9	ADB12064_12	Continuation (13 o
126	36	78.3	3517	8	ABZ42579	Abz42579 Human G p	c 199	35	76.1	113033	8	AAI54213	AaI54213 SR protei
127	36	78.3	3517	12	ADP21356	Adp21356 Gene GPR8	200	35	76.1	117328	13	ABD32886	AbD32886 Mouse can
128	36	78.3	3517	14	AEA47791	AeA47791 Nucleotid	c 201	35	76.1	160755	4	AAH88704	AaH88704 Human DNA
129	36	78.3	4146	3	AAC49442	AcA49442 Arabidops	c 202	35	76.1	337022	12	ADQ59416	AdQ59416 Human can
130	36	78.3	5106	12	ADG39768	Adg39768 Goose par	c 203	35	76.1	338780	14	ADZ13691	AdZ13691 Human can
131	36	78.3	9448	5	AAH41187	AaH41187 Human oli	c 204	35	76.1	339234	12	ADQ59437	AdQ59437 Human can
132	36	78.3	27082	4	AAK70447	AaK70447 Human imm	205	35	76.1	339234	14	ADZ13744	AdZ13744 Murine ca
133	36	78.3	77781	10	ADL15049	Adl15049 Human mel	c 206	35	76.1	349980	3	AAF21544	Aaf21544 Neisseria
134	36	78.3	80462	13	ABD33073	Abd33073 Murine ca	c 207	34	73.9	177	4	AAI23624	AaI23624 Probe #13
135	36	78.3	26187	11	ACN45182	Acn45182 Human gen	c 208	34	73.9	177	4	AAI21686	AaI21686 Probe #11
136	36	78.3	304326	13	AD512523	Ad512523 Rat senso	c 209	34	73.9	177	4	ABA76137	AbA76137 Human foe
137	35	76.1	100	8	ACD80688	AcD80688 E. coli K	c 210	34	73.9	177	4	ABA66764	AbA66764 Human foe
138	35	76.1	366	4	ACH4654	Ach4654 E. coli g	c 211	34	73.9	177	4	ABA68733	AbA68733 Human foe
139	35	76.1	434	9	ACH45995	Ach45995 Human inf	c 212	34	73.9	177	4	AAI46975	AaI46975 Probe #15
140	35	76.1	501	13	ACN49747	Acn49747 Cotton pr	c 213	34	73.9	177	4	AAI56794	AaI56794 Probe #25
141	35	76.1	570	4	AAI18144	AaI18144 Probe #80	c 214	34	73.9	177	4	AAI48935	AaI48935 Probe #17
142	35	76.1	570	4	ABA63113	AbA63113 Human foe	c 215	34	73.9	177	4	ABA48851	AbA48851 Human bre
143	35	76.1	570	4	AAI43151	AaI43151 Probe #11	c 216	34	73.9	177	4	ABA50766	AbA50766 Human bre
144	35	76.1	570	4	AAK11541	AaK11541 Human bra	c 217	34	73.9	177	4	ABA33828	AbA33828 Probe #12
145	35	76.1	570	4	AB536986	Ab536986 Human liv	c 218	34	73.9	177	4	ABA35697	AbA35697 Probe #14
146	35	76.1	601	13	ACN62852	Acn62852 Cotton ca	c 219	34	73.9	177	4	AAK40920	AaK40920 Human bon
147	35	76.1	639	10	ADC91934	Adc91934 E. faeciu	c 220	34	73.9	177	4	AAK42859	AaK42859 Human bon
148	35	76.1	724	6	ABL01444	AbL01444 Murine ap	c 221	34	73.9	177	4	AAK15195	AaK15195 Human bra
149	35	76.1	794	4	AAH70972	AaH70972 Human cer	c 222	34	73.9	177	4	AAK17076	AaK17076 Human bra
150	35	76.1	999	13	ADT20083	Adt20083 Plant cdn	c 223	34	73.9	177	4	ABS40501	AbS40501 Human liv

C 224	34	73.9	177	4	ABSS0394	AbS0394 Human liv	297	34	73.9	972	10	ADC92179	AdC92179 E. faeciu
C 225	34	73.9	177	4	ABSA4290	ABe4290 Human liv	298	34	73.9	1068	6	ABQ69274	ABq69274 Listeria
C 226	34	73.9	177	5	AAI07376	AAi07376 Probe #73	299	34	73.9	1119	6	ABQ67773	ABq67773 Listeria
C 227	34	73.9	177	5	AAI09239	AAi09239 Probe #92	300	34	73.9	1127	7	ABQ15696	ABq15696 Human pro
C 228	34	73.9	177	6	ABSI6914	ABs16914 Human gen	301	34	73.9	1163	6	ABN89009	ABn89009 Human pro
C 229	34	73.9	177	6	ABSI14876	ABs14876 Human gen	302	34	73.9	1176	2	AAQ38832	AAq38832 Pre-growt
C 230	34	73.9	177	6	ABSA4289	ABs4289 Human gen	C 303	34	73.9	1191	2	AAV37362	AAv37362 Streptoco
C 231	34	73.9	186	4	AAI27771	AAi27771 Probe #17	C 304	34	73.9	1242	8	ACA51261	ACa51261 Prokaryot
C 232	34	73.9	186	4	ABRA0638	ABa0638 Probe #19	C 305	34	73.9	1242	8	ACA52180	ACa52180 Prokaryot
C 233	34	73.9	186	4	AAKS0757	AAk0757 Human bon	306	34	73.9	1272	13	ADX82508	AdX82508 Leptospir
C 234	34	73.9	186	4	AAK24758	AAk24758 Human bra	307	34	73.9	1338	3	AAAC54805	AAa54805 Arabidops
C 235	34	73.9	186	4	ABSS0348	ABs0348 Human liv	C 308	34	73.9	1338	8	ACA27777	ACa27777 Prokaryot
C 236	34	73.9	186	6	ABSA2428	ABs2428 Human gen	309	34	73.9	1341	3	AAAC39265	AAa39265 Arabidops
C 237	34	73.9	264	10	ACP70724	ACp70724 Photorhab	C 310	34	73.9	1347	4	AAAS23419	AAa23419 Candida a
C 238	34	73.9	264	10	ACP70727	ACp70727 Photorhab	C 311	34	73.9	1347	6	ABZ31751	ABz31751 Candida a
C 239	34	73.9	280	10	ABX85549	ABx85549 Corn ear -	C 312	34	73.9	1352	8	ACA48629	ACa48629 Prokaryot
C 240	34	73.9	323	4	AAH71331	AAh71331 Human cer	C 313	34	73.9	1547	14	ADZ62433	ADz62433 Murine Tr
C 241	34	73.9	336	6	ABL78702	ABl78702 Human ova	C 314	34	73.9	1561	4	AAAS23809	AAa23809 Candida a
C 242	34	73.9	408	4	AAI14422	AAi14422 Probe #43	C 315	34	73.9	1566	8	ACA23012	ACa23012 Prokaryot
C 243	34	73.9	408	4	ABAS6147	ABa6147 Human foe	C 316	34	73.9	1576	6	AAAL42578	AAa42578 Human ald
C 244	34	73.9	408	4	AAI35795	AAi35795 Probe #44	C 317	34	73.9	1596	10	ADI21900	Adi21900 Novel hum
C 245	34	73.9	408	4	ABAA5644	ABa5644 Human bre	C 318	34	73.9	1624	13	ADX52401	AdX52401 Plant ful
C 246	34	73.9	408	4	ABAS25798	ABa25798 Probe #42	C 319	34	73.9	1722	13	ADRG1489	ADr61489 Cotton cd
C 247	34	73.9	408	4	AAK29831	AAk29831 Human bra	C 320	34	73.9	1729	13	ADX61419	AdX61419 Plant ful
C 248	34	73.9	408	4	AAK04337	AAk04337 Human bra	C 321	34	73.9	1737	4	AAAF94472	AAf94472 Human byd
C 249	34	73.9	408	4	ABAS9479	ABa9479 Human liv	C 322	34	73.9	1801	12	ADM47691	Adm47691 Polynucle
C 250	34	73.9	408	5	AAI04245	AAi04245 Probe #42	C 323	34	73.9	2000	6	ABZ16814	ABz16814 Arabidops
C 251	34	73.9	408	6	ABE04392	ABe04392 Human gen	C 324	34	73.9	2069	13	ADR07772	ADr07772 Full leng
C 252	34	73.9	423	9	ACH15715	ACH15715 Human adu	C 325	34	73.9	2083	8	ACC46090	ACC46090 Human dit
C 253	34	73.9	428	12	ADL10337	ADl10337 Cat flea	C 326	34	73.9	2368	13	ADS49263	ADs49263 Bacterial
C 254	34	73.9	435	3	AAAT9604	AAa9604 Pinus rad	C 327	34	73.9	2456	8	ABX34505	ABx34505 Human mdd
C 255	34	73.9	436	4	AAKS9973	AAk9973 Human imm	C 328	34	73.9	2456	10	ADE31312	AdE31312 Human dia
C 256	34	73.9	459	4	AAAL09655	AAa09655 Human bre	C 329	34	73.9	2563	4	AAH98872	AAh98872 Human EST
C 257	34	73.9	475	4	AAI12495	AAi12495 Probe #24	C 330	34	73.9	2595	8	ADA68592	ADa68592 Arabidops
C 258	34	73.9	475	4	ABAS4200	ABa54200 Human foe	C 331	34	73.9	2686	2	AAQ10163	AAq10163 Cyclomalt
C 259	34	73.9	475	4	AAI33850	AAi33850 Probe #25	C 332	34	73.9	2687	2	AAQ01792	AAq01792 Sequence
C 260	34	73.9	475	4	ABRA3747	ABa3747 Human bre	C 333	34	73.9	2795	10	ADA53183	AdA53183 Human cod
C 261	34	73.9	475	4	ABAS3951	ABa3951 Probe #24	C 334	34	73.9	2899	13	ADT14926	ADt14926 Plant CDN
C 262	34	73.9	475	4	AAK27916	AAk27916 Human bon	C 335	34	73.9	2996	2	AAQ12366	AAq12366 Gene enco
C 263	34	73.9	475	4	AAK02477	AAk02477 Human bra	C 336	34	73.9	3101	10	ADA52725	ADa52725 Human cod
C 264	34	73.9	475	4	ABAS27499	ABa27499 Human liv	C 337	34	73.9	3118	10	ADJ80184	ADj80184 Novel hum
C 265	34	73.9	475	5	AAI02405	AAi02405 Probe #23	C 338	34	73.9	3205	12	ADP04642	ADp04642 Sea squir
C 266	34	73.9	475	6	ABSO2371	ABs02371 Human gen	C 339	34	73.9	3310	4	AAH54779	AAh54779 S. epidr
C 267	34	73.9	481	6	ABL77892	ABl77892 Human ova	C 340	34	73.9	3370	2	AAQ79534	AAq79534 Bovine tr
C 268	34	73.9	510	6	ABK78460	ABk78460 Bacillus	C 341	34	73.9	3523	4	AAH55048	AAh55048 S. epidr
C 269	34	73.9	519	3	ABN81110	ABn81110 Shrimp po	C 342	34	73.9	3614	13	ADX49696	ADx49696 plant ful
C 270	34	73.9	542	6	ABN63218	ABn63218 Human can	C 343	34	73.9	4225	14	ABE21661	ABe21661 Indehiace
C 271	34	73.9	569	4	AAI18835	AAi18835 Probe #87	C 344	34	73.9	4267	13	ACN38342	ACn38342 Tumour-as
C 272	34	73.9	569	4	ABAS1014	ABa31014 Probe #94	C 345	34	73.9	4842	3	AAA37077	AAa37077 Human PRO
C 273	34	73.9	569	4	AAK38063	AAk38063 Human bon	C 346	34	73.9	4842	4	AAAF54343	AAf54343 DNA enco
C 274	34	73.9	569	4	AAK12344	AAk12344 Human bra	C 347	34	73.9	4842	4	AAAF46069	AAa46069 Human DNA
C 275	34	73.9	569	4	ABSI7682	ABs17682 Human liv	C 348	34	73.9	4842	8	ACA89519	ACa89519 cDNA enco
C 276	34	73.9	569	6	ABSI2070	ABs12070 Human gen	C 349	34	73.9	4842	8	ACA73529	ACa73529 Human sec
C 277	34	73.9	577	4	ABAS63915	ABa63915 Human foe	C 350	34	73.9	4842	8	ACA05844	ACa05844 Human sec
C 278	34	73.9	577	4	AAI44035	AAi44035 Probe #12	C 351	34	73.9	4842	8	ACA66678	ACa66678 cDNA enco
C 279	34	73.9	577	4	ABSI37754	ABs37754 Human liv	C 352	34	73.9	4842	8	ACF20253	ACf20253 Human sec
C 280	34	73.9	577	6	ABSI12158	ABs12158 Human gen	C 353	34	73.9	4842	8	ACF19639	ACf19639 Human sec
C 281	34	73.9	633	10	ADK54918	ADk54918 Plant DNA	C 354	34	73.9	4842	8	ACD21927	ACd21927 Human sec
C 282	34	73.9	635	4	AAAF24640	AAf24640 Fragment	C 355	34	73.9	4842	8	ACF13092	ACf13092 Human sec
C 283	34	73.9	692	5	ADL63121	ADl63121 Human ova	C 356	34	73.9	4842	8	ACD25195	ACd25195 Human sec
C 284	34	73.9	699	6	ABN92146	ABn92146 Staphyloc	C 357	34	73.9	4842	8	ACF00244	ACf00244 Human sec
C 285	34	73.9	699	13	ADSO1841	ADs01841 Staphyloc	C 358	34	73.9	4842	8	ACD72301	ACd72301 Novel hum
C 286	34	73.9	765	10	ABX07587	ABx07587 S. pneumo	C 359	34	73.9	4842	8	ACD04825	ACd04825 Novel hum
C 287	34	73.9	768	8	ACA50092	ACa50092 Prokaryot	C 360	34	73.9	4842	8	ACD18286	ACd18286 Human sec
C 288	34	73.9	768	13	ADK43862	ADk43862 Streptoco	C 361	34	73.9	4842	8	ACD08293	ACd08293 Human sec
C 289	34	73.9	771	6	ABN89008	ABn89008 Human pro	C 362	34	73.9	4842	8	ACA88727	ACa88727 Novel hum
C 290	34	73.9	786	4	AAAF4462	AAf4462 Human hyd	C 363	34	73.9	4842	8	ACA70169	ACa70169 Human sec
C 291	34	73.9	798	4	AAH76199	AAh76199 Human dru	C 364	34	73.9	4842	8	ACD12391	ACd12391 Novel hum
C 292	34	73.9	798	13	ADR91761	ADr91761 Novel S.	C 365	34	73.9	4842	8	ACC74306	ACC74306 Human sec
C 293	34	73.9	798	11	AAAS5631	AAa5631 Streptoco	C 366	34	73.9	4842	8	ACD15934	ACd15934 Human sec
C 294	34	73.9	852	11	ACN79855	ACn79855 Breast ca	C 367	34	73.9	4842	8	ACD25502	ACd25502 Novel hum
C 295	34	73.9	858	10	ADCS9084	ADc9084 Novel hum	C 368	34	73.9	4842	8	ACD17979	ACd17979 Human sec
C 296	34	73.9	913	11	ACN84923	ACn84923 Breast ca	C 369	34	73.9	4842	8	ACC88266	ACC88266 Human sec

370	34	73.9	4842	8	ACD21620	Human sec	443	34	73.9	4842	9	ACA99120	Novel hum
371	34	73.9	4842	8	ACD18687	Human sec	444	34	73.9	4842	9	ACC91752	Human sec
372	34	73.9	4842	8	ABX98297	Human cDN	445	34	73.9	4842	9	ACD11163	Novel hum
373	34	73.9	4842	8	ACD14048	Human PRO	446	34	73.9	4842	9	ACD15013	Human sec
374	34	73.9	4842	8	ACD09828	Human sec	447	34	73.9	4842	9	ACD11777	Human sec
375	34	73.9	4842	8	ACC88573	Human sec	448	34	73.9	4842	9	ACC95906	Human sec
376	34	73.9	4842	8	ACD21313	Human sec	449	34	73.9	4842	9	ACF16469	Human sec
377	34	73.9	4842	8	ABX75685	Human cDN	450	34	73.9	4842	9	ACF02587	Human sec
378	34	73.9	4842	8	ABX97888	Human PRO	451	34	73.9	4842	9	ACF02894	Human sec
379	34	73.9	4842	8	ACA97364	Novel hum	452	34	73.9	4842	9	ACF21481	Human sec
380	34	73.9	4842	8	ACA57827	Human PRO	453	34	73.9	4842	9	ACF10165	Human sec
381	34	73.9	4842	8	ACD14355	Human PRO	454	34	73.9	4842	9	ACF78058	Human sec
382	34	73.9	4842	8	ACC91138	Human sec	455	34	73.9	4842	9	ACF46763	Human sec
383	34	73.9	4842	8	ACC88880	Human sec	456	34	73.9	4842	9	ACD49526	Human sec
384	34	73.9	4842	8	ACD07077	Human PRO	457	34	73.9	4842	9	ACF28293	Human sec
385	34	73.9	4842	8	ACA67528	Human PRO	458	34	73.9	4842	9	ACD88983	Human sec
386	34	73.9	4842	8	ACC81583	Human sec	459	34	73.9	4842	9	ACD84378	Human PRO
387	34	73.9	4842	8	ACC89187	Human sec	460	34	73.9	4842	9	ACD99152	CDNA enco
388	34	73.9	4842	8	ACC86543	Human sec	461	34	73.9	4842	9	ADA78041	Human sec
389	34	73.9	4842	8	ACC89801	Human sec	462	34	73.9	4842	9	ACF48894	Human sec
390	34	73.9	4842	8	ACC92380	Human sec	463	34	73.9	4842	9	ACD09214	Human sec
391	34	73.9	4842	8	ACA72608	Human PRO	464	34	73.9	4842	9	ACF12007	Human sec
392	34	73.9	4842	8	ACA89126	Human PRO	465	34	73.9	4842	9	ACF41241	Human sec
393	34	73.9	4842	8	ACA69862	Human sec	466	34	73.9	4842	9	ACF15855	Human sec
394	34	73.9	4842	8	ACA97005	Novel hum	467	34	73.9	4842	9	ACF16162	Human sec
395	34	73.9	4842	8	ACA91001	Novel hum	468	34	73.9	4842	9	ACD31989	Human sec
396	34	73.9	4842	8	ACA70783	Human sec	469	34	73.9	4842	9	ACF18797	Human sec
397	34	73.9	4842	8	ACA95293	Novel hum	470	34	73.9	4842	9	ACF09244	Human sec
398	34	73.9	4842	8	ACC86236	Human sec	471	34	73.9	4842	9	ACF78365	Human sec
399	34	73.9	4842	8	ACC90108	Human sec	472	34	73.9	4842	9	ACF51964	Human sec
400	34	73.9	4842	8	ACD12716	Human sec	473	34	73.9	4842	9	ACF26451	Human sec
401	34	73.9	4842	8	ACF19946	Human sec	474	34	73.9	4842	9	ACF24244	Human sec
402	34	73.9	4842	8	ABX76890	Human PRO	475	34	73.9	4842	9	ACF63555	Human sec
403	34	73.9	4842	8	ACA73222	Novel hum	476	34	73.9	4842	9	ACF50429	Human sec
404	34	73.9	4842	8	ACA68765	Novel hum	477	34	73.9	4842	9	ACH07900	Human sec
405	34	73.9	4842	8	ACA74609	CDNA enco	478	34	73.9	4842	9	ACF13706	Human sec
406	34	73.9	4842	8	ACA70476	Human sec	479	34	73.9	4842	9	ACD41632	Human sec
407	34	73.9	4842	8	ACD14662	Human PRO	480	34	73.9	4842	9	ACF32045	Human sec
408	34	73.9	4842	8	ACA68334	Novel hum	481	34	73.9	4842	9	ACF23323	Human sec
409	34	73.9	4842	8	ABX98799	Novel hum	482	34	73.9	4842	9	ACF40013	Human sec
410	34	73.9	4842	8	ACC81376	Human sec	483	34	73.9	4842	9	ACD45535	Human sec
411	34	73.9	4842	8	ACA95600	Novel hum	484	34	73.9	4842	9	ACF53192	Human sec
412	34	73.9	4842	8	ACD04518	Novel hum	485	34	73.9	4842	9	ACF27372	Human sec
413	34	73.9	4842	8	ACC87959	Human sec	486	34	73.9	4842	9	ACF45210	Human sec
414	34	73.9	4842	8	ACF12621	Human sec	487	34	73.9	4842	9	ACF29828	Human sec
415	34	73.9	4842	8	ACA96336	Human PRO	488	34	73.9	4842	9	ACD89904	Human sec
416	34	73.9	4842	8	ACA65110	Human PRO	489	34	73.9	4842	9	ACD84685	Human sec
417	34	73.9	4842	8	ACA73836	Human sec	490	34	73.9	4842	9	ACF98845	Human sec
418	34	73.9	4842	8	ACA74248	Novel hum	491	34	73.9	4842	9	ACF77137	Human sec
419	34	73.9	4842	8	ACA96643	Human PRO	492	34	73.9	4842	9	ACF49815	Human sec
420	34	73.9	4842	8	ACD10749	CDNA enco	493	34	73.9	4842	9	ACF49815	Human sec
421	34	73.9	4842	8	ACC91445	Human sec	494	34	73.9	4842	9	ACF50122	Human sec
422	34	73.9	4842	8	ACD02780	CDNA enco	495	34	73.9	4842	9	ACD09521	Human sec
423	34	73.9	4842	8	ACC87345	Human sec	496	34	73.9	4842	9	ACD08600	Human sec
424	34	73.9	4842	8	ACC85929	Human sec	497	34	73.9	4842	9	ACF12314	Human sec
425	34	73.9	4842	8	ACA65417	Human PRO	498	34	73.9	4842	9	ACC94822	Human sec
426	34	73.9	4842	8	ACA94234	Human sec	499	34	73.9	4842	9	ACD22541	Human sec
427	34	73.9	4842	8	ACA97978	Human PRO	500	34	73.9	4842	9	ACF15241	Human sec
428	34	73.9	4842	8	ACA91480	Novel hum	501	34	73.9	4842	9	ACC97336	Human sec
429	34	73.9	4842	8	ACA90694	Novel hum	502	34	73.9	4842	9	ACC92366	Human sec
430	34	73.9	4842	8	ACD16241	Human sec	503	34	73.9	4842	9	ACF14013	Human sec
431	34	73.9	4842	8	ACD17402	Human sec	504	34	73.9	4842	9	ACF14320	Human sec
432	34	73.9	4842	8	ACC92059	Human sec	505	34	73.9	4842	9	ACF09551	Human sec
433	34	73.9	4842	8	ACA74916	CDNA enco	506	34	73.9	4842	9	ACD68380	Novel hum
434	34	73.9	4842	8	ACA91787	Human PRO	507	34	73.9	4842	9	ACD45842	Human sec
435	34	73.9	4842	8	ACA71431	Human sec	508	34	73.9	4842	9	ACD47991	Human sec
436	34	73.9	4842	8	ACC90831	Human sec	509	34	73.9	4842	9	ACD67722	CDNA enco
437	34	73.9	4842	8	ACA65841	CDNA enco	510	34	73.9	4842	9	ACF25530	Human sec
438	34	73.9	4842	8	ACA94986	CDNA enco	511	34	73.9	4842	9	ACF29214	Human sec
439	34	73.9	4842	8	ACD16548	Human sec	512	34	73.9	4842	9	ACD84992	Human sec
440	34	73.9	4842	8	ACD15627	Human sec	513	34	73.9	4842	9	ACD84071	Human PRO
441	34	73.9	4842	8	ABX16730	Human cDN	514	34	73.9	4842	9	ACD88062	Human sec
442	34	73.9	4842	9	ACA97671	Human PRO	515	34	73.9	4842	9	ACF30749	Human sec

516	34	73.9	4842	9	ACF32352	Human sec	589	34	73.9	4842	9	ACF13399	Human sec
517	34	73.9	4842	9	ACH12012	cdNA enco	590	34	73.9	4842	9	ACF03201	Human sec
518	34	73.9	4842	9	ACH12319	cdNA enco	591	34	73.9	4842	9	ACF78672	Human sec
519	34	73.9	4842	9	ACH12319	cdNA enco	592	34	73.9	4842	9	ACF11393	Human sec
520	34	73.9	4842	9	ACH12319	cdNA enco	593	34	73.9	4842	9	ACF50736	Human sec
521	34	73.9	4842	9	ACH12319	cdNA enco	594	34	73.9	4842	9	ACF34231	Human sec
522	34	73.9	4842	9	ACH12319	cdNA enco	595	34	73.9	4842	9	ACD46456	Human sec
523	34	73.9	4842	9	ACH12319	cdNA enco	596	34	73.9	4842	9	ACD48298	Human sec
524	34	73.9	4842	9	ACH12319	cdNA enco	597	34	73.9	4842	9	ACF27679	Human sec
525	34	73.9	4842	9	ACH12319	cdNA enco	598	34	73.9	4842	9	ACF24551	Human sec
526	34	73.9	4842	9	ACH12319	cdNA enco	599	34	73.9	4842	9	ACD85606	Human sec
527	34	73.9	4842	9	ACH12319	cdNA enco	600	34	73.9	4842	9	ACD90211	Human sec
528	34	73.9	4842	9	ACH12319	cdNA enco	601	34	73.9	4842	9	ACD83764	Human PRO
529	34	73.9	4842	9	ACH12319	cdNA enco	602	34	73.9	4842	9	ACF49201	Human sec
530	34	73.9	4842	9	ACH12319	cdNA enco	603	34	73.9	4842	9	ACH07286	Human sec
531	34	73.9	4842	9	ACH12319	cdNA enco	604	34	73.9	4842	9	ACH07593	Human sec
532	34	73.9	4842	9	ACH12319	cdNA enco	605	34	73.9	4842	9	ACH08207	Human sec
533	34	73.9	4842	9	ACH12319	cdNA enco	606	34	73.9	4842	9	ACH11398	Human sec
534	34	73.9	4842	9	ACH12319	cdNA enco	607	34	73.9	4842	9	ACH11705	cdNA enco
535	34	73.9	4842	9	ACH12319	cdNA enco	608	34	73.9	4842	9	ACH10356	Human sec
536	34	73.9	4842	9	ACH12319	cdNA enco	609	34	73.9	4842	9	ACH01359	Human sec
537	34	73.9	4842	9	ACH12319	cdNA enco	610	34	73.9	4842	9	ACH40934	Human sec
538	34	73.9	4842	9	ACH12319	cdNA enco	611	34	73.9	4842	9	ACH24274	Human sec
539	34	73.9	4842	9	ACH12319	cdNA enco	612	34	73.9	4842	9	ACH31375	Human sec
540	34	73.9	4842	9	ACH12319	cdNA enco	613	34	73.9	4842	9	ACH17876	Human sec
541	34	73.9	4842	9	ACH12319	cdNA enco	614	34	73.9	4842	9	ACH32659	Human sec
542	34	73.9	4842	9	ACH12319	cdNA enco	615	34	73.9	4842	9	ACH40320	Human sec
543	34	73.9	4842	9	ACH12319	cdNA enco	616	34	73.9	4842	9	ACH48280	Human sec
544	34	73.9	4842	9	ACH12319	cdNA enco	617	34	73.9	4842	9	ACH38229	Human sec
545	34	73.9	4842	9	ACH12319	cdNA enco	618	34	73.9	4842	9	ACH25165	Human sec
546	34	73.9	4842	9	ACH12319	cdNA enco	619	34	73.9	4842	9	ACH27065	Human sec
547	34	73.9	4842	9	ACH12319	cdNA enco	620	34	73.9	4842	9	ACH29521	Human sec
548	34	73.9	4842	9	ACH12319	cdNA enco	621	34	73.9	4842	9	ACH87755	Human sec
549	34	73.9	4842	9	ACH12319	cdNA enco	622	34	73.9	4842	9	ACH76216	Human sec
550	34	73.9	4842	9	ACH12319	cdNA enco	623	34	73.9	4842	9	ACH49508	Human sec
551	34	73.9	4842	9	ACH12319	cdNA enco	624	34	73.9	4842	9	ACH43965	Human sec
552	34	73.9	4842	9	ACH12319	cdNA enco	625	34	73.9	4842	9	ACH06310	cdNA enco
553	34	73.9	4842	9	ACH12319	cdNA enco	626	34	73.9	4842	9	ACH06617	cdNA enco
554	34	73.9	4842	9	ACH12319	cdNA enco	627	34	73.9	4842	9	ACH83358	Human sec
555	34	73.9	4842	9	ACH12319	cdNA enco	628	34	73.9	4842	9	ACH92673	Human sec
556	34	73.9	4842	9	ACH12319	cdNA enco	629	34	73.9	4842	9	ACH93287	Human sec
557	34	73.9	4842	9	ACH12319	cdNA enco	630	34	73.9	4842	9	ACH19332	Human sec
558	34	73.9	4842	9	ACH12319	cdNA enco	631	34	73.9	4842	9	ACH13023	Human sec
559	34	73.9	4842	9	ACH12319	cdNA enco	632	34	73.9	4842	9	ACH06481	Human sec
560	34	73.9	4842	9	ACH12319	cdNA enco	633	34	73.9	4842	9	ACH94515	Human sec
561	34	73.9	4842	9	ACH12319	cdNA enco	634	34	73.9	4842	9	ACH97943	Human sec
562	34	73.9	4842	9	ACH12319	cdNA enco	635	34	73.9	4842	9	ACH94208	Human sec
563	34	73.9	4842	9	ACH12319	cdNA enco	636	34	73.9	4842	9	ACH42162	Human sec
564	34	73.9	4842	9	ACH12319	cdNA enco	637	34	73.9	4842	9	ACH31068	Human sec
565	34	73.9	4842	9	ACH12319	cdNA enco	638	34	73.9	4842	9	ACH43097	cdNA enco
566	34	73.9	4842	9	ACH12319	cdNA enco	639	34	73.9	4842	9	ACH43404	cdNA enco
567	34	73.9	4842	9	ACH12319	cdNA enco	640	34	73.9	4842	9	ACH14934	Human sec
568	34	73.9	4842	9	ACH12319	cdNA enco	641	34	73.9	4842	9	ACH01666	Human sec
569	34	73.9	4842	9	ACH12319	cdNA enco	642	34	73.9	4842	9	ACH31738	Human sec
570	34	73.9	4842	9	ACH12319	cdNA enco	643	34	73.9	4842	9	ACH67415	cdNA enco
571	34	73.9	4842	9	ACH12319	cdNA enco	644	34	73.9	4842	9	ACH48605	Human sec
572	34	73.9	4842	9	ACH12319	cdNA enco	645	34	73.9	4842	9	ACH48912	Human sec
573	34	73.9	4842	9	ACH12319	cdNA enco	646	34	73.9	4842	9	ACH51350	Human sec
574	34	73.9	4842	9	ACH12319	cdNA enco	647	34	73.9	4842	9	ACH54113	Human sec
575	34	73.9	4842	9	ACH12319	cdNA enco	648	34	73.9	4842	9	ACH25837	Human sec
576	34	73.9	4842	9	ACH12319	cdNA enco	649	34	73.9	4842	9	ACH39150	Human sec
577	34	73.9	4842	9	ACH12319	cdNA enco	650	34	73.9	4842	9	ACH28907	Human sec
578	34	73.9	4842	9	ACH12319	cdNA enco	651	34	73.9	4842	9	ACH90824	Human sec
579	34	73.9	4842	9	ACH12319	cdNA enco	652	34	73.9	4842	9	ACH86527	Human sec
580	34	73.9	4842	9	ACH12319	cdNA enco	653	34	73.9	4842	9	ACH05389	cdNA enco
581	34	73.9	4842	9	ACH12319	cdNA enco	654	34	73.9	4842	9	ACH65185	Human sec
582	34	73.9	4842	9	ACH12319	cdNA enco	655	34	73.9	4842	9	ACH20401	Human sec
583	34	73.9	4842	9	ACH12319	cdNA enco	656	34	73.9	4842	9	ACH43658	Human sec
584	34	73.9	4842	9	ACH12319	cdNA enco	657	34	73.9	4842	9	ACH09128	Human sec
585	34	73.9	4842	9	ACH12319	cdNA enco	658	34	73.9	4842	9	ACH09435	Human sec
586	34	73.9	4842	9	ACH12319	cdNA enco	659	34	73.9	4842	9	ACH78653	Human sec
587	34	73.9	4842	9	ACH12319	cdNA enco	660	34	73.9	4842	9	ACH09858	Human sec
588	34	73.9	4842	9	ACH12319	cdNA enco	661	34	73.9	4842	9	ACH51043	Human sec

662	34	73.9	4842	9	ACF23937	Human sec
663	34	73.9	4842	9	ACD88369	Human sec
664	34	73.9	4842	9	ACH09742	Human sec
665	34	73.9	4842	9	ACH10663	Human sec
666	34	73.9	4842	9	ACH11470	Human sec
667	34	73.9	4842	9	ACC96520	Human sec
668	34	73.9	4842	9	ACH04482	Human CDN
669	34	73.9	4842	9	ACC98550	Human sec
670	34	73.9	4842	9	ACF41855	Human sec
671	34	73.9	4842	9	ACF16776	Human sec
672	34	73.9	4842	9	ACD32296	Human sec
673	34	73.9	4842	9	ACD30454	Human sec
674	34	73.9	4842	9	ACD41325	Human sec
675	34	73.9	4842	9	ACF07709	Human sec
676	34	73.9	4842	9	ACF31124	Human sec
677	34	73.9	4842	9	ACF77444	Human sec
678	34	73.9	4842	9	ACF11086	Human sec
679	34	73.9	4842	9	ACF32966	Human sec
680	34	73.9	4842	9	ACF26144	Human sec
681	34	73.9	4842	9	ACD83457	Human PRO
682	34	73.9	4842	9	ACF23630	Human sec
683	34	73.9	4842	9	ACF43044	Human sec
684	34	73.9	4842	9	ACF43351	Human sec
685	34	73.9	4842	9	ACH06003	CDNA enco
686	34	73.9	4842	9	ACH08821	Human sec
687	34	73.9	4842	9	ACC90415	Human sec
688	34	73.9	4842	9	ACF10779	Human sec
689	34	73.9	4842	9	ACC93594	Human sec
690	34	73.9	4842	9	ACC96213	Human sec
691	34	73.9	4842	9	ACD24888	Human sec
692	34	73.9	4842	9	ACF01973	Human sec
693	34	73.9	4842	9	ACF22095	Human sec
694	34	73.9	4842	9	ACF22709	Human sec
695	34	73.9	4842	9	ACF08937	Human sec
696	34	73.9	4842	9	ACF33273	Human sec
697	34	73.9	4842	9	ACF45727	Human sec
698	34	73.9	4842	9	ACF48587	Human sec
699	34	73.9	4842	9	ACD47377	Human sec
700	34	73.9	4842	9	ACD49219	Human sec
701	34	73.9	4842	9	ACF37922	Human sec
702	34	73.9	4842	9	ACF30135	Human sec
703	34	73.9	4842	9	ACD87448	Human sec
704	34	73.9	4842	9	ACF62037	Human sec
705	34	73.9	4842	9	ACH10970	Human sec
706	34	73.9	4842	9	ACD10135	Human sec
707	34	73.9	4842	9	ACD16860	CDNA enco
708	34	73.9	4842	9	ACC99157	Human sec
709	34	73.9	4842	9	ACF00551	Human sec
710	34	73.9	4842	9	ACD41018	Human sec
711	34	73.9	4842	9	ACF14627	Human sec
712	34	73.9	4842	9	ACF22402	Human sec
713	34	73.9	4842	9	ACF78979	Human sec
714	34	73.9	4842	9	ACD68026	Novel hum
715	34	73.9	4842	9	ACF11700	Human sec
716	34	73.9	4842	9	ACF51657	Human sec
717	34	73.9	4842	9	ACF33580	Human sec
718	34	73.9	4842	9	ACD49833	Human sec
719	34	73.9	4842	9	ACF37615	Human sec
720	34	73.9	4842	9	ACF28600	Human sec
721	34	73.9	4842	9	ACD88676	Human sec
722	34	73.9	4842	9	ACF75295	Human sec
723	34	73.9	4842	9	ACF61116	Human sec
724	34	73.9	4842	9	ACF44272	Human sec
725	34	73.9	4842	9	ACH08514	Human sec
726	34	73.9	4842	9	ACC93901	Human sec
727	34	73.9	4842	9	ACD21006	Human sec
728	34	73.9	4842	9	ACF06788	Human sec
729	34	73.9	4842	9	ACD20699	Human sec
730	34	73.9	4842	9	ACD22848	Human sec
731	34	73.9	4842	9	ACF41548	Human sec
732	34	73.9	4842	9	ACF07095	Human sec
733	34	73.9	4842	9	ACF77751	Human sec
734	34	73.9	4842	9	ACD46149	Human sec
735	34	73.9	4842	9	ACF23937	Human sec
736	34	73.9	4842	9	ACD88369	Human sec
737	34	73.9	4842	9	ACH09742	Human sec
738	34	73.9	4842	9	ACH10663	Human sec
739	34	73.9	4842	9	ACH11470	Human sec
740	34	73.9	4842	9	ACC96520	Human sec
741	34	73.9	4842	9	ACH04482	Human CDN
742	34	73.9	4842	9	ACC98550	Human sec
743	34	73.9	4842	9	ACF41855	Human sec
744	34	73.9	4842	9	ACF16776	Human sec
745	34	73.9	4842	9	ACD32296	Human sec
746	34	73.9	4842	9	ACD30454	Human sec
747	34	73.9	4842	9	ACD41325	Human sec
748	34	73.9	4842	9	ACF07709	Human sec
749	34	73.9	4842	9	ACF31124	Human sec
750	34	73.9	4842	9	ACF77444	Human sec
751	34	73.9	4842	9	ACF11086	Human sec
752	34	73.9	4842	9	ACF32966	Human sec
753	34	73.9	4842	9	ACF26144	Human sec
754	34	73.9	4842	9	ACD83457	Human PRO
755	34	73.9	4842	9	ACF23630	Human sec
756	34	73.9	4842	9	ACF43044	Human sec
757	34	73.9	4842	9	ACF43351	Human sec
758	34	73.9	4842	9	ACH06003	CDNA enco
759	34	73.9	4842	9	ACH08821	Human sec
760	34	73.9	4842	9	ACC90415	Human sec
761	34	73.9	4842	9	ACF10779	Human sec
762	34	73.9	4842	9	ACC93594	Human sec
763	34	73.9	4842	9	ACC96213	Human sec
764	34	73.9	4842	9	ACD24888	Human sec
765	34	73.9	4842	9	ACF01973	Human sec
766	34	73.9	4842	9	ACF22095	Human sec
767	34	73.9	4842	9	ACF22709	Human sec
768	34	73.9	4842	9	ACF08937	Human sec
769	34	73.9	4842	9	ACF33273	Human sec
770	34	73.9	4842	9	ACF45727	Human sec
771	34	73.9	4842	9	ACF48587	Human sec
772	34	73.9	4842	9	ACD47377	Human sec
773	34	73.9	4842	9	ACD49219	Human sec
774	34	73.9	4842	9	ACF37922	Human sec
775	34	73.9	4842	9	ACF30135	Human sec
776	34	73.9	4842	9	ACD87448	Human sec
777	34	73.9	4842	9	ACF62037	Human sec
778	34	73.9	4842	9	ACH10970	Human sec
779	34	73.9	4842	9	ACD10135	Human sec
780	34	73.9	4842	9	ACD16860	CDNA enco
781	34	73.9	4842	9	ACC99157	Human sec
782	34	73.9	4842	9	ACF00551	Human sec
783	34	73.9	4842	9	ACD41018	Human sec
784	34	73.9	4842	9	ACF14627	Human sec
785	34	73.9	4842	9	ACF22402	Human sec
786	34	73.9	4842	9	ACF78979	Human sec
787	34	73.9	4842	9	ACD68026	Novel hum
788	34	73.9	4842	9	ACF11700	Human sec
789	34	73.9	4842	9	ACF51657	Human sec
790	34	73.9	4842	9	ACF33580	Human sec
791	34	73.9	4842	9	ACD49833	Human sec
792	34	73.9	4842	9	ACF37615	Human sec
793	34	73.9	4842	9	ACF28600	Human sec
794	34	73.9	4842	9	ACD88676	Human sec
795	34	73.9	4842	9	ACF75295	Human sec
796	34	73.9	4842	9	ACF61116	Human sec
797	34	73.9	4842	9	ACF44272	Human sec
798	34	73.9	4842	9	ACH08514	Human sec
799	34	73.9	4842	9	ACC93901	Human sec
800	34	73.9	4842	9	ACD21006	Human sec
801	34	73.9	4842	9	ACF06788	Human sec
802	34	73.9	4842	9	ACD20699	Human sec
803	34	73.9	4842	9	ACD22848	Human sec
804	34	73.9	4842	9	ACF41548	Human sec
805	34	73.9	4842	9	ACF07095	Human sec
806	34	73.9	4842	9	ACF77751	Human sec
807	34	73.9	4842	9	ACD46149	Human sec
808	34	73.9	4842	9	ACF23937	Human sec
809	34	73.9	4842	9	ACD88369	Human sec
810	34	73.9	4842	9	ACH09742	Human sec
811	34	73.9	4842	9	ACH10663	Human sec
812	34	73.9	4842	9	ACH11470	Human sec
813	34	73.9	4842	9	ACC96520	Human sec
814	34	73.9	4842	9	ACH04482	Human CDN
815	34	73.9	4842	9	ACC98550	Human sec
816	34	73.9	4842	9	ACF41855	Human sec
817	34	73.9	4842	9	ACF16776	Human sec
818	34	73.9	4842	9	ACD32296	Human sec
819	34	73.9	4842	9	ACD30454	Human sec
820	34	73.9	4842	9	ACD41325	Human sec
821	34	73.9	4842	9	ACF07709	Human sec
822	34	73.9	4842	9	ACF31124	Human sec
823	34	73.9	4842	9	ACF77444	Human sec
824	34	73.9	4842	9	ACF11086	Human sec
825	34	73.9	4842	9	ACF32966	Human sec
826	34	73.9	4842	9	ACF26144	Human sec
827	34	73.9	4842	9	ACD83457	Human PRO
828	34	73.9	4842	9	ACF23630	Human sec
829	34	73.9	4842	9	ACF43044	Human sec
830	34	73.9	4842	9	ACF43351	Human sec
831	34	73.9	4842	9	ACH06003	CDNA enco
832	34	73.9	4842	9	ACH08821	Human sec
833	34	73.9	4842	9	ACC90415	Human sec
834	34	73.9	4842	9	ACF10779	Human sec
835	34	73.9	4842	9	ACC93594	Human sec
836	34	73.9	4842	9	ACC96213	Human sec
837	34	73.9	4842	9	ACD24888	Human sec
838	34	73.9	4842	9	ACF01973	Human sec
839	34	73.9	4842	9	ACF22095	Human sec
840	34	73.9	4842	9	ACF22709	Human sec
841	34	73.9	4842	9	ACF08937	Human sec
842	34	73.9	4842	9	ACF33273	Human sec
843	34	73.9	4842	9	ACF45727	Human sec
844	34	73.9	4842	9	ACF48587	Human sec
845	34	73.9	4842	9	ACD47377	Human sec
846	34	73.9	4842	9	ACD49219	Human sec
847	34	73.9	4842	9	ACF37922	Human sec
848	34	73.9	4842	9	ACF30135	Human sec
849	34	73.9	4842	9	ACD87448	Human sec
850	34	73.9	4842	9	ACF62037	Human sec
851	34	73.9	4842	9	ACH10970	Human sec
852	34	73.9	4842	9	ACD10135	Human sec
853	34	73.9	4842	9	ACD16860	CDNA enco
854	34	73.9	4842	9	ACC99157	Human sec
855	34	73.9	4842	9	ACF00551	Human sec
856	34	73.9	4842	9	ACD41018	Human sec
857	34	73.9	4842	9	ACF14627	Human sec
858	34	73.9	4842	9	ACF22402	Human sec
859	34	73.9	4842	9		

808	34	73.9	4842	12	ADH04039	Human cDN
809	34	73.9	4842	12	ADH03562	Human cDN
810	34	73.9	4842	12	ADH26219	Novel hum
811	34	73.9	4842	12	ADH33188	Human PRO
812	34	73.9	4842	12	ADH04516	Human cDN
813	34	73.9	4842	12	ADH61517	Human cDN
814	34	73.9	4842	12	ADJ54927	Human PRO
815	34	73.9	4842	12	ADJ64698	Human PRO
816	34	73.9	4842	12	ADN31594	Novel hum
817	34	73.9	4842	12	ADM36641	Novel hum
818	34	73.9	4842	12	ADM40446	Novel hum
819	34	73.9	4842	12	ADL94716	Human cDN
820	34	73.9	4842	12	ADN38054	Human cDN
821	34	73.9	4842	14	ADN85137	Human cDN
822	34	73.9	5000	4	AA814506	Human GST
823	34	73.9	5015	2	AA89163	Human sli
824	34	73.9	5015	2	AA819946	Human sli
825	34	73.9	5015	6	ABL92110	Human Tum
826	34	73.9	5015	8	ABZ34808	Coding se
827	34	73.9	5015	10	ABX72035	DNA encod
828	34	73.9	5015	12	ADQ17593	Human sof
829	34	73.9	5176	2	AA61026	Human sli
830	34	73.9	6524	6	AA150068	Bovine di
831	34	73.9	6703	2	AAV49536	Adenylate
832	34	73.9	7924	6	ABK40070	Human che
833	34	73.9	7924	6	ABL34139	Human imm
834	34	73.9	7951	4	ABL10710	Drosophil
835	34	73.9	9603	12	ADQ22279	Human sof
836	34	73.9	10357	2	AAV52324	Streptoco
837	34	73.9	12706	4	ABL19664	Drosophil
838	34	73.9	13382	14	ADW44488	Zebrafish
839	34	73.9	14899	4	ABL20584	Drosophil
840	34	73.9	16001	14	AAE62857	Rat arc-c
841	34	73.9	17294	6	ABL32986	Human imm
842	34	73.9	19026	10	ADC86778	Human GPC
843	34	73.9	19387	5	AA34561	Human DNA
844	34	73.9	27781	10	ABX77186	Genomic D
845	34	73.9	29220	4	AA827653	DNA encod
846	34	73.9	29220	4	AA827652	DNA encod
847	34	73.9	29220	10	ADB94455	Novel hum
848	34	73.9	29220	10	ADB94456	Novel hum
849	34	73.9	32207	5	ABA19666	Human ner
850	34	73.9	33352	9	ADA02846	Human FGF
851	34	73.9	33352	10	ADB72584	Human FGF
852	34	73.9	33352	10	ADC85325	Mouse Fgf
853	34	73.9	33352	12	ADW74441	Human car
854	34	73.9	35558	14	AD212667	Human can
855	34	73.9	38142	13	ABD32682	Mouse can
856	34	73.9	50000	10	ADC51644	BmPV gen
857	34	73.9	52211	11	ACN44892	Mouse gen
858	34	73.9	60381	11	ACN44494	Human gen
859	34	73.9	65854	4	AAK86282	Human imm
860	34	73.9	76798	6	ABN97454	Gene #395
861	34	73.9	76798	14	ADX06902	Cyclin-de
862	34	73.9	81905	6	ABQ69244	Listeria
863	34	73.9	82689	6	ABQ67198	Listeria
864	34	73.9	91352	12	ADN94799	DNA encod
865	34	73.9	110000	2	AAT42063	(12 o
866	34	73.9	110000	2	AA20248	(7 of
867	34	73.9	110000	6	ABA03041	(4 of
868	34	73.9	110000	10	ADF77343	(11 o
869	34	73.9	110000	10	ABS56454	(17 o
870	34	73.9	110000	10	ACP67367	(39 o
871	34	73.9	110000	10	ACP65388	(10 o
872	34	73.9	110000	10	ACP65388	(11 o
873	34	73.9	110000	12	ADQ34435	(6 of
874	34	73.9	110000	13	ABD32966	(3 of
875	34	73.9	110218	11	ACN44744	Mouse gen
876	34	73.9	111309	2	AA202050	Borrelia
877	34	73.9	113585	12	ADJ19197	Human int
878	34	73.9	152330	11	ACN45070	Human gen
879	34	73.9	172637	6	ABN83124	Human vol
880	34	73.9	172637	14	ABE80195	Human tra
881	34	73.9	237961	6	ABQ80552	Human Can
882	34	73.9	260209	6	AB556564	Human SUL
883	34	73.9	260209	12	ADN16204	Human sul
884	34	73.9	335913	5	AAI61371	Soybean 2
885	34	73.9	335913	5	AAI61372	Soybean 2
886	34	73.9	349901	10	ADC86940	Human GPC
887	34	73.9	349938	10	ADC87621	Human GPC
888	34	73.9	349980	13	ADT05649	Haemophil
889	33.5	72.8	43179	13	ABD33316	Murine ca
890	33	71.7	46	8	ABX12827	PCR prime
891	33	71.7	46	8	ABX12826	PCR prime
892	33	71.7	160	6	AB572467	Human gen
893	33	71.7	237	6	ABL36943	Human col
894	33	71.7	268	3	AAA45903	Human sec
895	33	71.7	276	6	ABN77667	Human ORF
896	33	71.7	317	3	AAA79371	Human ORF
897	33	71.7	341	2	AA113870	Eucalyptu
898	33	71.7	341	6	AB599665	Enterococ
899	33	71.7	342	11	ACH94847	Klebsiell
900	33	71.7	358	6	ABV96038	Human pan
901	33	71.7	365	14	ADY78551	Human cDN
902	33	71.7	391	3	AACT75788	Human ORF
903	33	71.7	407	8	ABX35766	Bovine ES
904	33	71.7	420	9	ACH50167	Human leu
905	33	71.7	423	9	ACH17234	Human adu
906	33	71.7	434	6	ABN79166	Human ORF
907	33	71.7	453	2	AAAT19901	Human gen
908	33	71.7	473	9	ACH27626	Human adu
909	33	71.7	474	13	ADX48626	Plant ful
910	33	71.7	478	9	ACH23537	Human adu
911	33	71.7	501	10	ADA48859	Rat gene
912	33	71.7	503	12	ADN12789	Human pro
913	33	71.7	504	9	ACH36471	Human end
914	33	71.7	553	14	ADY66149	S. mansoni
915	33	71.7	555	5	ADL43769	Human ova
916	33	71.7	555	13	ACN50810	Corton an
917	33	71.7	561	6	ABT10514	Human bre
918	33	71.7	561	13	ADQ51788	Novel can
919	33	71.7	573	13	ADQ52106	Novel can
920	33	71.7	576	8	ACA40109	Prokaryot
921	33	71.7	578	6	ABN65776	Human can
922	33	71.7	584	4	ABA60729	Human foe
923	33	71.7	584	4	AAI40620	Probe #93
924	33	71.7	584	4	ABA28800	Probe #72
925	33	71.7	584	4	AAK34904	Human bon
926	33	71.7	584	4	AAK09012	Human bra
927	33	71.7	584	4	AB34660	Human liv
928	33	71.7	584	6	AB509423	Human gen
929	33	71.7	586	12	ADQ18367	Human sof
930	33	71.7	588	5	ABV54666	Human pro
931	33	71.7	618	10	ABZ39207	N. gonorr
932	33	71.7	619	4	AA522848	Human cDN
933	33	71.7	654	5	AA594238	DNA encod
934	33	71.7	661	14	ACL65370	M. xanthu
935	33	71.7	664	13	ADQ51672	Novel can
936	33	71.7	664	4	AA522612	Human cDN
937	33	71.7	667	11	ACN86219	Breast ca
938	33	71.7	669	11	ACN83325	Breast ca
939	33	71.7	693	13	AD59190	Bacterial
940	33	71.7	713	10	ABX94439	Rice endo
941	33	71.7	784	10	AD333703	Mouse mit
942	33	71.7	805	11	ACN85050	Breast ca
943	33	71.7	814	6	AB565033	Invertebr
944	33	71.7	814	8	ABX10355	DNA encod
945	33	71.7	838	4	AAH07013	Human cDN
946	33	71.7	871	11	ACN84319	Breast ca
947	33	71.7	984	8	ACA35557	Prokaryot
948	33	71.7	987	8	ACA32645	Prokaryot
949	33	71.7	987	8	ACA48994	Prokaryot
950	33	71.7	987	8	ACA51460	Prokaryot
951	33	71.7	1001	3	AAH51453	Human UGT
952	33	71.7	1005	11	ACH96404	Klebsiell
953	33	71.7	1108	13	ADX27769	Plant ful

954 33 71.7 1125 11 AC97334
 955 33 71.7 1137 8 ACA28118
 956 33 71.7 1158 8 ACA28064
 957 33 71.7 1176 10 ADF00794
 958 33 71.7 1233 6 ABN68764
 959 33 71.7 1248 5 AAS78645
 960 33 71.7 1260 4 ABL11645
 961 33 71.7 1311 13 AD845407
 962 33 71.7 1389 2 AAQ70737
 963 33 71.7 1451 4 ABL19703
 964 33 71.7 1512 6 ABQ70530
 965 33 71.7 1526 13 ADX14741
 966 33 71.7 1551 13 ADT16569
 967 33 71.7 1578 2 AAT42223
 968 33 71.7 1578 2 AAT79597
 969 33 71.7 1581 6 ABZ13389
 970 33 71.7 1593 2 AAT58563
 971 33 71.7 1593 2 AAT58562
 972 33 71.7 1648 4 AAH13852
 973 33 71.7 1707 4 AAH17937
 974 33 71.7 1814 6 AAD34303
 975 33 71.7 1819 12 ADI16268
 976 33 71.7 1824 6 AAS16434
 977 33 71.7 1836 12 ADM36206
 978 33 71.7 1849 13 ADX13354
 979 33 71.7 1922 5 ADL63169
 980 33 71.7 1924 3 AAC49880
 981 33 71.7 1928 3 AAC40588
 982 33 71.7 1981 13 ADO81502
 983 33 71.7 2000 6 ABZ16711
 984 33 71.7 2000 11 ACL37157
 985 33 71.7 2000 11 ACL36459
 986 33 71.7 2004 7 ADZ74729
 987 33 71.7 2134 13 ADX51629
 988 33 71.7 2167 5 AAC82727
 989 33 71.7 2175 13 ADX50311
 990 33 71.7 2222 13 ADT17735
 991 33 71.7 2232 6 ABZ12933
 992 33 71.7 2242 12 ADJ75712
 993 33 71.7 2242 14 ADZ62274
 994 33 71.7 2286 10 ACF71229
 995 33 71.7 2338 4 AAH16583
 996 33 71.7 2528 6 ABK35880
 997 33 71.7 2591 13 ADX59577
 998 33 71.7 2633 10 ADD47986
 999 33 71.7 2676 10 ADC37206
 c1000 33 71.7 2750 10 ACC60692

ALIGNMENTS

RESULT 1
 ID ABX40605
 AC ABX40605 standard; cDNA; 343 BP.

20-FEB-2003 (first entry)

Bovine EST associated with lactation/muscle/fat deposition #5770.

Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 muscle deposition; fat deposition; genome mapping; gene identification;
 gene analysis; cattle breeding.

Bos Taurus.

US2002137139-A1.

26-SEP-2002.

24-SEP-2001; 2001US-00960352.

PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.

XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.

XX Claim 2; SEQ ID NO 5770; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived from
 CC cattle, and the LMPD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridization between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMPD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139

SQ Sequence 343 BP; 40 A; 146 C; 108 G; 49 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.34 Length: 343
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x ABX40605 (1-343)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 Db 173 AACCTCTTCTCAGCGGCAACACAGCTG 199

RESULT 2

ACLS6146
 ID ACL56146 standard; cDNA; 505 BP.

XX AC ACL56146;

XX 24-MAR-2005 (first entry)

XX Human colon cancer differentially expressed polynucleotide, SEQ ID:2281.

XX Differential expression; diagnosis; therapy; drug screening; cancer;
 KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;

```
KW ss.
XX Homo sapiens.
OS
XX WO200500087-A2.
FN
XX
XX 06-JAN-2005.
XX
XX 13-MAY-2004; 2004WO-US015421.
XX
XX 03-JUN-2003; 2003US-0475872P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Randazzo P, Moler B, Escobedo J, Garcia PD;
XX WPI; 2005-075421/08.
XX
XX New isolated polynucleotides, which are differentially expressed in colon
XX cancer cell, useful for treating cancer, e.g. colon cancer, breast
XX cancer, or pancreatic cancer.
XX
XX Claim 1; SEQ ID NO 2281; 97pp; English.
XX
XX The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
XX are differentially expressed in colon cancer cells. The invention also
XX relates to vectors and host cells comprising a differentially expressed
XX polynucleotide of the invention; a method for detecting a cancerous cell
XX by detection of a gene product of the polynucleotides; a method for
XX inhibiting a cancerous phenotype of a cell by inhibiting a gene product
XX of the polynucleotides; a method of treating an individual with cancer by
XX administration of a modulator of a gene product of the polynucleotides;
XX and an isolated antibody that specifically binds to a polypeptide encoded
XX by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
XX antibodies, and methods are useful for the detection of cancerous cells;
XX for the diagnosis, prognosis and management of cancer; for the
XX identification of agents that modulate the phenotype of cancerous cells;
XX for the identification of therapeutic targets for cancer chemotherapy;
XX and for the treatment of cancer, especially colon cancer and metastasized
XX colon cancer, but also breast or pancreatic cancer. The polynucleotides
XX are also useful as a source of probes or primers for use in diagnostic
XX methods. The differentially expressed polynucleotides or their encoded
XX proteins can additionally be used as vaccines to modulate primary immune
XX responses for the prevention or treatment of cancer. The present sequence
XX represents a specifically claimed polynucleotide which is differentially
XX expressed in colon cancer. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 505 BP; 68 A; 202 C; 149 G; 85 T; 0 U; 1 Other;
XX
XX Alignment Scores:
XX Pred. No.: 2.1 Length: 505
XX Score: 46.00 Matches: 9
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 100.0% Indels: 0
XX DB: 14 Gaps: 0
XX
XX US-10-774-176-12 (1-9) x ACL56146 (1-505)
XX
XX QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
XX
XX Db 174 AACCTCTCTTACCGCACACAGCTG 200
XX
XX RESULT 3
XX ABT07721
XX ID ABT07721 standard; DNA; 927 BP.
XX
XX AC ABT07721;
XX
XX DT 14-NOV-2002 (first entry)
XX
```

```
XX Breast cancer-associated gene sequence 29.
XX
XX Gene; ds; breast cancer; breast cancer-associated gene sequence;
XX drug development; pharmacogenetics; biosensor development.
XX
XX Unidentified.
XX
XX WO200259377-A2.
XX
XX 01-AUG-2002.
XX
XX 24-JAN-2002; 2002WO-US002242.
XX
XX 24-JAN-2001; 2001US-0263965P.
XX
XX 02-FEB-2001; 2001US-0285928P.
XX
XX 09-APR-2001; 2001US-00829472.
XX
XX 04-APR-2001; 2001US-0282698P.
XX
XX 04-MAY-2001; 2001US-0288590P.
XX
XX 29-MAY-2001; 2001US-0294443P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Mack DH, Gish KC, Afar D;
XX WPI; 2002-583738/62.
XX
XX N-PSDB; ABJ05564.
XX
XX Detecting a breast cancer-associated transcript in a patient's cell,
XX useful for diagnosing breast cancer, comprises contacting a biological
XX sample with a polynucleotide that selectively hybridizes with breast
XX cancer nucleic acids.
XX
XX Claim 9; Page 372; 414pp; English.
XX
XX The invention comprises a method of detecting a breast cancer-associated
XX transcript in a cell from a patient. The method of the invention involves
XX contacting a biological sample from the patient with a nucleotide that
XX hybridizes to one of the 69 breast cancer-associated gene sequences shown
XX in the specification. The method of the invention is useful in the
XX diagnosis or prognosis of breast cancer, and for detecting genes that are
XX up or down-regulated in breast cancer cells. Genes identified by the
XX method of the invention can be used in diagnostic purposes and also as
XX targets for screening for therapeutic compounds that modulate breast
XX cancer (e.g. hormones or antibodies). Identification of genes that are
XX over or under expressed in breast cancer can additionally provide high-
XX resolution, high-sensitivity datasets which can be used in the areas of
XX diagnostics, therapeutics, drug development, pharmacogenetics, protein
XX structure and biosensor development. DNA sequences ABT07693 - ABT07761
XX represent the 69 breast cancer-associated gene sequences of the invention
XX
XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 4.26 Length: 927
XX Score: 46.00 Matches: 9
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 100.0% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-10-774-176-12 (1-9) x ABT07721 (1-927)
XX
XX QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
XX
XX Db 283 AACCTCTCTTACCGCACACAGCTG 309
XX
XX RESULT 4
XX ABX76333
XX ID ABX76333 standard; DNA; 927 BP.
XX
XX AC ABX76333;
XX
XX
```

DT 02-APR-2003 (first entry)
 XX Lung cancer-associated polynucleotide #197.
 XX
 XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 XX
 XX WO200286443-A2.
 PN 31-OCT-2002.
 XX
 XX 18-APR-2002; 2002WO-US012476.
 XX
 XX 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0324370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Aziz N, Murray R;
 XX
 XX WPI; 2003-093161/08.
 DR P-PSDB; ABUS6604.
 XX
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 XX Claim 22; Page 336; 453pp; English.
 XX
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.26 Length: 927
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-12 (1-9) x ABX76333 (1-927)
 QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 ||||||||||||||||||||||||||||

Db 283 AACCTCTTCTTACCGGCAACCAGCTG 309
 RESULT 5
 ADB80503
 ID ADB80503 standard; DNA; 927 BP.
 XX
 AC ADB80503;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 XX Ovarian cancer-associated transcript #34.
 XX
 XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; ds; gene.
 KW
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH CDS 1..927
 FT /*tag= a
 XX
 XX WO2002102235-A2.
 PN 27-DEC-2002.
 XX
 PD 18-JUN-2002; 2002WO-US019297.
 XX
 PF 18-JUN-2001; 2001US-0299234P.
 XX
 PR 27-AUG-2001; 2001US-0315287P.
 PR 05-SEP-2001; 2001US-0317544P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Mack DH, Gish KC;
 FI
 XX WPI; 2003-167431/16.
 DR P-PSDB; ADB80504.
 XX
 XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.
 XX
 XX Claim 10; Page 297; 332pp; English.
 PS
 XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.26 Length: 927
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-12 (1-9) x ADB80503 (1-927)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 Db 283 AACCTCTCTTACCGCACACGCTG 309

RESULT 6
 ADN38723
 ID ADN38723 standard; cDNA; 927 BP.
 AC ADN38723;
 XX
 XX 17-JUN-2004 (first entry)
 XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
 DE
 XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX
 XX Homo sapiens.
 XX
 XX W02003042661-A2.
 XX
 XX 22-MAY-2003.
 XX
 XX 13-NOV-2002; 2002WO-US036810.
 XX
 XX 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-03353394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0352250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 XX (BOSB-) BOS BIOTECHNOLOGY INC.
 XX
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 XX WPI; 2003-468649/44.
 DR P-PSDB; ADN38724.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 XX Claim 8; SEQ ID NO 41; 1395pp; English.
 XX
 XX The invention relates to nucleic acids and proteins (ADN38693-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;

and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.

XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.26 Length: 927
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

US-10-774-176-12 (1-9) x ADN38723 (1-927)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 Db 283 AACCTCTCTTACCGCACACGCTG 309

RESULT 7
 ABV99349
 ID ABV99349 standard; DNA; 1156 BP.
 XX
 XX AC ABV99349;
 XX
 XX 27-JAN-2003 (first entry)
 DT
 XX Human NOV8a coding sequence.
 DE
 XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
 KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
 XX
 XX Homo sapiens.
 OS
 XX W0200272771-A2.
 FN
 XX
 XX 19-SEP-2002.
 PD
 XX
 XX 08-MAR-2002; 2002WO-US007288.
 XX
 XX 08-MAR-2001; 2001US-0274101P.
 PR 08-MAR-2001; 2001US-0274194P.
 PR 08-MAR-2001; 2001US-0274281P.
 PR 08-MAR-2001; 2001US-0274322P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275578P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 16-MAR-2001; 2001US-0276776P.
 PR 19-MAR-2001; 2001US-0276994P.
 PR 20-MAR-2001; 2001US-0277239P.
 PR 20-MAR-2001; 2001US-0277321P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.

PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280333P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-032530P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 21-NOV-2001; 2001US-0333272P.
PR 03-DEC-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 04-DEC-2001; 2001US-0338092P.
PR 03-JAN-2002; 2001US-0337185P.
PR 08-MAR-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX (CURA-) CURAGEN CORP.
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
XX Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
XX Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
XX Voss EZ, Malyankar UM, Anderson DM, Patturajan M, Miller CE;
XX Traupler RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
XX Zhong M;
XX WPI; 2002-732824/79.
DR P-PSDB; ABP70071.
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders e.g. Diabetes, cancer,
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
PT disorders, and asthma.
XX Claim 16; Page 114-115; 619pp; English.

XX The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's

CC disease, immune disorders, hematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods

XX SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores: 5.5 Length: 1156
Pred. No.: 46.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 6 Gaps: 0
DB:

US-10-774-176-12 (1-9) x ABV99349 (1-1156)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 223 AACCTCTTCTTACCGGCAACACGCTG 249
 |||||

RESULT 8

ABK87175

ID ABK87175 standard; cDNA; 1260 BP.

XX AC ABK87175;

XX DT 07-OCT-2002 (first entry)

XX DE cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.

XX KW Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
XX KW cell proliferative disorder; infection; inflammatory condition;
XX KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
XX KW foetal abnormality; foetal sex determination; gene; ss.

XX OS Felis sp.

XX FH Key Location/Qualifiers

XX FT CDS 1..1260

XX FT /*tag= a

XX FT /product= "5T4 protein"

XX FN WO200238612-A2.

XX PD 16-MAY-2002.

XX XX 13-NOV-2001; 2001WO-GB005004.

XX XX 13-NOV-2000; 2000WO-GB004317.

XX XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX FA Myers K, Drury N, Carroll M;

XX PI WPI; 2002-557449/59.

XX DR P-PSDB; AU98694.

XX XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the

XX PT polypeptide, useful in preparation of vaccine for treating and/or

XX PT preventing cancer in a subject, preferably a dog or cat.

XX XX Claim 4; Page 68; 68pp; English.

XX CC The present invention relates to the isolation of canine and feline

XX CC oncofoetal leucine-rich glycoproteins known as 5T4, and the

XX CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in

CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes feline 5T4 protein

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 6.08 Length: 1260
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x ABK87175 (1-1260)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 DB 280 AACCTCTCTCCACCGGCATCAGCTG 306

RESULT 9
 ADB97513

ID ADB97513 standard; DNA; 1260 BP.

XX AC ADB97513;

XX DT 04-DEC-2003 (first entry)

XX DE Feline 5T4 antigen DNA.

XX KW Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.

XX OS Unidentified.

XX PH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"

XX PN WO2003068816-A1.

XX PD 21-AUG-2003.

XX PF 13-FEB-2003; 2003WO-GB000670.

XX PR 13-FEB-2002; 2002GB-00003419.

XX XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX XX Carroll M, Kingsman S, Redchenko I;

XX XX WPI; 2003-637141/60.

XX DR P-PSDB; ADB97520.

XX PT New major histocompatibility complex class I peptide epitopes from human
 PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.

XX PS Disclosure; Page 67; 73pp; English.

XX

CC The invention relates to a novel Major Histocompatibility Complex (MHC)
 CC class I peptide epitope of the 5T4 antigen. The invention further
 CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
 CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
 CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 5T4 epitope and/its encoding nucleic acid; a method
 CC comprising detecting the presence of the 5T4 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 5T4 epitope in conjunction with an MHC class
 CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC particularly cancer. The detection method is useful for diagnosing or
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline
 CC 5T4 antigen coding DNA of the invention.

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 6.08 Length: 1260
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-12 (1-9) x ADB97513 (1-1260)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 DB 280 AACCTCTCTCCACCGGCATCAGCTG 306

RESULT 10
 ADB97452

ID ADB97452 standard; DNA; 1260 BP.

XX AC ADB97452;

XX DT 04-DEC-2003 (first entry)

XX DE DNA encoding feline 5T4 protein.

XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
 KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.

XX OS Unidentified.

XX PH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"

XX PN WO2003068815-A2.

XX PD 21-AUG-2003.

XX PF 13-FEB-2003; 2003WO-GB000618.

XX PR 13-FEB-2002; 2002GB-00003420.

XX XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX XX Carroll M, Harrop R, Kingsman S;

XX WPI; 2003-663795/62.

XX DR P-PSDB; ADB97455.

XX New Major Histocompatibility Complex class II peptide epitope of 5T4,
PT useful for manufacturing a medicament for diagnosing, preventing and/or
PT treating a disease, e.g. cancer.
XX
XX Disclosure, Page 49; 63pp; English.
XX
XX The invention relates to a Major Histocompatibility Complex (MHC) class
CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
CC clone has a cytostatic activity, as it is useful in manufacturing a
CC medicament for preventing and/or treating a disease, particularly cancer.
CC The methods are useful for detecting T-cells capable of specifically
CC recognising a peptide epitope in conjunction with an MHC molecule, for
CC diagnosing or monitoring the progression of a cancerous disease, or for
CC detecting the presence of a peptide or nucleic acid using an agent. The
CC MHC class II peptide epitope of the invention can be used in gene therapy
CC or as part of a vaccine. This polynucleotide sequence represents the DNA
CC coding for the feline 5T4 protein.
XX
SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores: Length: 1260
Pred. No.: 6.08 Matches: 9
Score: 46.00 Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 10

US-10-774-176-12 (1-9) x ADB97452 (1-1260)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 ||||| ||||| ||||| |||||
Db 280 AACCTCTTCTCACCAGCAATCAGCTG 306

RESULT 11
AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX
AC AAA27058;
XX
DT 22-AUG-2000 (first entry)
XX
DE Human 5T4 tumour-associated antigen gene.
XX
XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
XX Homo sapiens.
XX
XX WO200029428-A2.
PN
XX
PD 25-MAY-2000.
XX
XX 18-NOV-1999; 99WO-GB003859.
PF
XX
XX 18-NOV-1998; 98GB-00025303.
PR
XX 27-JAN-1999; 99GB-00001739.
PR
XX 30-JUL-1999; 99GB-00017995.
PR
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX
XX Carroll MW, Myers KA;
PI
XX
XX WPI; 2000-387735/33.
DR
XX
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
PT response useful in vaccinating against and in treating tumors.
PT
XX
XX Example 2; Page 78; 79pp; English.
XX
XX The present sequence encodes the human 5T4 tumour-associated antigen

CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
CC carcinomas but has a highly restricted expression pattern in normal adult
CC tissues. It appears to be strongly correlated to metastasis in colorectal
CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
CC induced were inoculated with a virus expression vector containing the
CC present sequence. The 5T4 antigen was shown to be effective at eliciting
CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
CC the antigen and the antigen itself can be used to elicit an immune
CC response, preferably CTL or an antibody response in a subject
XX
SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores: Length: 1263
Pred. No.: 6.1 Matches: 9
Score: 46.00 Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 3

US-10-774-176-12 (1-9) x AAA27058 (1-1263)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 ||||| ||||| ||||| |||||
Db 283 AACCTCTTCTTACCGCAACACGCTG 309

RESULT 12
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX
AC AAF89736;
XX
DT 23-JUL-2001 (first entry)
XX
XX Nucleotide sequence of canine 5T4 protein.
XX
XX Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
KW hypersensitivity; autoimmune disease; central nervous system disorder;
KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
KW Helicobacter-related disease; immune disorder; ss.
XX
XX Canis sp.
XX
XX Key Location/Qualifiers
FT CDS 1..1263
FT /*tag= a
FT /product= "5T4"
XX
XX WO200136486-A2.
PN
XX
XX 25-MAY-2001.
PD
XX
XX 13-NOV-2000; 2000WO-GB004317.
PF
XX
XX 18-NOV-1999; 99WO-GB003859.
PR
XX 15-FEB-2000; 2000GB-00003527.
PR
XX 02-MAR-2000; 2000GB-00005071.
PR
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX
XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard PM;
PI Myers KA;
XX
XX WPI; 2001-343805/36.
DR
XX P-PSDB; AAB83839.
DR
XX
XX Use of single chain antibody capable of recognizing a disease associated
PT molecule for manufacturing a medicament for preventing and/or treating a
PT disease condition associated with disease associated molecule.
PT
XX
XX Disclosure; Fig 26; 118pp; English.
PS

XX The specification describes the use of a single chain antibody (ScFv),
CC which is capable of recognizing a disease associated molecule in the
CC manufacture of a medicament for the prevention and treatment of a disease
CC condition. The ScFv antibody is useful in the manufacture of a
CC medicament, for affecting a disease in vivo, for preparing a
CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
CC treatment of a disease. The ScFv antibody is also useful for treating
CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
CC diseases, cancers, central nervous system disorders including Parkinson's
CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
CC related diseases, and other immune disorders. The present sequence
CC encodes a 574 protein, which is used to produce ScFv of the invention
XX
SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6.1 Length: 1263
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0
US-10-774-176-12 (1-9) x AAF89736 (1-1263)
QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
DB 283 AACCTCTTCTCTCAGCGGCACACGCTG 309
RESULT 13
ID ABK87174 standard; cDNA; 1263 BP.
XX
AC ABK87174;
DT 07-OCT-2002 (first entry)
XX
XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 574.
DE Canine; dog; oncofoetal leucine-rich glycoprotein; 574; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX
OS Canis sp.
XX
FH Key Location/Qualifiers
FT CDS 1..1263
FT /*tag= a
FT /product= "574 protein"
XX
XX WO200238612-A2.
PN
XX
XX 16-MAY-2002.
PD
XX
XX 13-NOV-2001; 2001WO-GB005004.
PF
XX
XX 13-NOV-2000; 2000WO-GB004317.
PR
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX
XX Myers K, Drury N, Carroll M;
PI
XX
XX WPI; 2002-557449/59.
DR
XX
XX P-PSDB; AAU98693.
DR
XX
XX Novel canine or feline 574 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
PS Claim 1; Page 67; 68pp; English.

XX The present invention relates to the isolation of canine and feline
CC oncofoetal leucine-rich glycoproteins known as 574, and the
CC polynucleotide sequences encoding them. The 574 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 574 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 574
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus (es).
XX
XX The present sequence encodes canine 574 protein
SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6.1 Length: 1263
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-12 (1-9) x ABK87174 (1-1263)
QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
DB 283 AACCTCTTCTCTCAGCGGCACACGCTG 309
RESULT 14
ID AAD56199 standard; DNA; 1331 BP.
XX
AC AAD56199;
DT 07-AUG-2003 (first entry)
XX
DE Human LRRCAPS related DNA #6.
XX
KW Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
FN WO2003035831-A2.
XX
XX 01-MAY-2003.
PD
XX
XX 21-OCT-2002; 2002WO-US033540.
PF
XX
XX 22-OCT-2001; 2001US-0338733P.
PR
XX
XX 15-FEB-2002; 2002US-0357600P.
PR
XX
XX 01-MAR-2002; 2002US-0361196P.
PR
XX
XX (EXEL-) EXELIXIS INC.
PA
XX
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
PI
XX
XX WPI; 2003-421410/39.
DR
XX
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
PS Disclosure; Page 75-76; 99pp; English.

CC The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 6.48 Length: 1331
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 8 Indels: 0
DB: Gaps: 0

US-10-774-176-12 (1-9) x RAD56199 (1-1331)

QY 1 AsnLeuPheLeuThrGlyAenGlnLeu 9

DB 313 AACCTCTTCTTACCGGCAACACGCTG 339

RESULT 15

ADJ56299
ID ADJ56299 standard; cDNA; 2020 BP.

XX AC ADJ56299;

XX DT 06-MAY-2004 (first entry)

XX DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.

XX KW human; differential expression; transactivator; proto-oncogene;
XX KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
XX MYCN activated cell.

XX OS Homo sapiens.

XX PN US2003119009-A1.

XX PD 26-JUN-2003.

XX PF 25-FEB-2002; 2002US-00084817.

XX PR 23-FEB-2001; 2001US-0270784P.

XX PA (STUA/) STUART S G.

XX PA (NUCH/) NUCHTERN J G.

XX PA (PLOW/) PLOW S E.

XX PA (SHOH/) SHOHET J M.

XX PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;

XX WPI; 2003-635698/60.

XX PT New genes regulated by MYCN activation, useful in gene therapy,
XX particularly for treating a subject with e.g. neuroblastoma or other
XX cancers, or for diagnosing, staging or monitoring the treatment of the
XX cancer.

XX PS Claim 1; SEQ ID NO 105; 27pp; English.

XX CC This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
XX polynucleotide sequences that exhibit differential expression patterns in
XX cells activated by the transactivator MYCN, where MYCN is a proto-
XX oncogene that is amplified in neuroblastoma cells and is common in small
XX cell lung cancers. The present invention describes these cDNA molecules

CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.

SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 10.5 Length: 2020
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-774-176-12 (1-9) x ADJ56299 (1-2020)

QY 1 AsnLeuPheLeuThrGlyAenGlnLeu 9

DB 353 AACCTCTTCTTACCGGCAACACGCTG 379

RESULT 16

ACC51052
ID ACC51052 standard; cDNA; 2053 BP.

XX AC ACC51052;

XX DT 12-JUN-2003 (first entry)

XX DE Human bladder cancer associated cDNA sequence SEQ ID NO:192.

XX KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX PN WO2003003906-A2.

XX PD 16-JAN-2003.

XX PF 03-JUL-2002; 2002WO-US021338.

XX PR 03-JUL-2001; 2001US-0302814P.

XX PR 03-AUG-2001; 2001US-0310099P.

XX PR 08-NOV-2001; 2001US-0343705P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Mack DH, Aziz N;

XX WPI; 2003-201532/19.

XX DR P-PSDB; ABR48236.

XX PT Detecting a bladder cancer-associated transcript in a cell from a
XX patient, comprises contacting a biological sample from the patient with a
XX bladder cancer-associated polynucleotide or antibody.

XX PS Claim 6; Page 296; 307pp; English.

XX CC The present invention describes a method for detecting a bladder cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with a polynucleotide
XX that selectively hybridises to a sequence that is 80 % identical to a
XX table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
XX encode the human bladder cancer-associated proteins given in ABR48146 to

CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores: 10.7 Length: 2053
Pred. No.: 46.00 Matches: 9
Score: 46.00
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x ACC51052 (1-2053)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||
Db 367 AACCTCTTCCTTACCGGCACACGCTG 393

RESULT 17

ID ABX76332
ID ABX76332 standard; DNA; 2053 BP.

AC ABX76332;

XX 02-APR-2003 (first entry)

DT Lung cancer-associated polynucleotide #196.

DE Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antitasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

XX 10-MAY-2001; 2001US-0290492P.

XX 09-NOV-2001; 2001US-0339245P.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0334370P.

XX 12-APR-2002; 2002US-0372246P.

XX (BOSB-) BOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX P-PSDB; ABUS6603.

XX Detecting a lung cancer-associated transcript in a cell from a patient

XX for treating lung cancer, by contacting a biological sample from the

XX patient with a polynucleotide that exhibits increased or decreased

XX expression in lung cancer.

XX Claim 22; Page 335; 453pp; English.

CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 10.7 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x ABX76332 (1-2053)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||
Db 367 AACCTCTTCCTTACCGGCACACGCTG 393

RESULT 18

AAD56197

ID AAD56197 standard; DNA; 2053 BP.

XX AAD56197;

XX 07-AUG-2003 (first entry)

XX Human LRRCAPS DNA #11.

XX Human; p53 pathway; Leucine rich repeat capricious related protein;

XX LRRCAPS; cancer; gene therapy; ds.

XX Homo sapiens.

XX WO2003035831-A2.

XX 01-MAY-2003.

XX 21-OCT-2002; 2002WO-US033540.

XX 22-OCT-2001; 2001US-0338733P.

XX 15-FEB-2002; 2002US-0357600P.

XX 01-MAR-2002; 2002US-0361196P.

XX (EXEL-) EXELIXIS INC.

XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;

XX Francis-Lang H, Friedman L;

XX WPI; 2003-421410/39.

XX Identifying a candidate p53 pathway-modulating agent for treating cancer
XX comprises contacting an assay system comprising a purified leucine rich
XX repeat, capricious related polypeptide or nucleic acid with a test agent.

XX Example 5; Page 73-74; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway

PT Determining the presence or absence of a pathological cell in a patient, PT useful for diagnosing, prognosing or treating cancer, comprises detecting

PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO 39; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 10.7 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-12 (1-9) x ADN38721 (1-2053)
QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||
Db 367 AACCTCTTCCTTACCGCAACACGCTG 393

RESULT 21
ADL06473
ID ADL06473 standard; cDNA; 2053 BP.
XX
AC ADL06473;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human tumour-associated antigenic target (TAT) cDNA sequence #53.
XX
XX Human; tumour-associated antigenic target; TAT; cell death; tumour;
XX
XX cancer; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2004016225-A2.
XX
XX 26-FEB-2004.
XX
XX 19-AUG-2003; 2003WO-US025892.
XX
XX 19-AUG-2002; 2002US-0404809P.
XX
XX 21-AUG-2002; 2002US-0405645P.
XX
XX 23-SEP-2002; 2002US-0413192P.
XX
XX 15-OCT-2002; 2002US-0419008P.
XX
XX 15-NOV-2002; 2002US-0426847P.
XX
XX 02-JUL-2003; 2003US-0484959P.
XX
XX (GETH) GENENTECH INC.
XX
XX Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
XX
XX Spencer SD, Wu TD, Zhang Z;
XX
XX WPI; 2004-257144/24.
XX
XX P-PSDB; ADL06552.

XX New antibody that binds to a tumor-associated antigenic target (TAT) polypeptide, useful for preparing a composition for diagnosing or treating cancer.
XX
PS Claim 1; SEQ ID NO 53; 319pp; English.
XX
CC The present invention relates to the isolation of human tumour-associated antigenic target (TAT) polynucleotide and polypeptide sequences. Also disclosed is an antibody that binds to a TAT polypeptide. The antibody is a monoclonal antibody, an antibody fragment, a chimeric antibody or a humanized antibody. It is conjugated to a growth inhibitory agent. It is produced in bacteria or in CHO cells and induces death of a cell to which it binds. The antibody is useful for preparing a composition for diagnosing or treating tumours and cancer. The present sequence represents a human TAT cDNA sequence of the invention.
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 10.7 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-12 (1-9) x ADL06473 (1-2053)
QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||
Db 367 AACCTCTTCCTTACCGCAACACGCTG 393

RESULT 22
ADN03961
ID ADN03961 standard; cDNA; 2053 BP.
XX
AC ADN03961;
XX
XX 01-JUL-2004 (first entry)
XX
XX Antipsoriatic cDNA sequence #180.
XX
XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
XX Homo sapiens.
XX
XX WO2004028479-A2.
XX
XX 08-APR-2004.
XX
XX 25-SEP-2003; 2003WO-US030907.
XX
XX 25-SEP-2002; 2002US-0414006P.
XX
XX (GETH) GENENTECH INC.
XX
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX
XX Wu TD;
XX
XX WPI; 2004-305105/28.
XX
XX P-PSDB; ADN03962.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a pharmaceutical composition for diagnosing or treating psoriasis in a mammal.
XX
XX Claim 1; SEQ ID NO 355; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for treating psoriasis or a sequence having at least 80% identity to the above sequences. The nucleic acid is useful for preparing a composition for diagnosing or treating psoriasis in a mammal. This sequence

CC corresponds to one of the polynucleotides of the invention.
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 10.7 Length: 2053
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-12 (1-9) x ADN03961 (1-2053)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 DB 367 AACCTCTTCCTTACCGGCAACACGCTG 393

RESULT 23
 ADR25444
 ID ADR25444 standard; DNA; 2053 BP.

XX AC ADR25444;
 XX DT 21-OCT-2004 (first entry)

XX DE Breast cancer prognosis marker #1305.

XX KW ds; breast cancer; prognosis; gene expression; diagnosis.

XX OS Homo sapiens.

XX PN WO2004065545-A2.

XX PD 05-AUG-2004.

XX PF 15-JAN-2004; 2004WO-US001100.

XX PR 15-JAN-2003; 2003US-00342887.

XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
 XX PA (NECA-) NETHERLANDS CANCER INST.

XX PI Van't Veer LJ, He Y;

XX WPI; 2004-593473/57.

XX Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.

XX Disclosure; SEQ ID NO 1305; 226pp; English.

XX The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 10.7 Length: 2053
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-12 (1-9) x ADR25444 (1-2053)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 DB 367 AACCTCTTCCTTACCGGCAACACGCTG 393

RESULT 24

ACN38510

ID ACN38510 standard; cDNA; 2053 BP.

XX AC ACN38510;

XX DT 18-NOV-2004 (first entry)

XX DE Tumour-associated antigenic target (TAT) CDNA DNA103471, SEQ ID NO:2070.

XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;

XX KW tumour; diagnosis; cell proliferative disorder; breast cancer;

XX KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;

XX KW central nervous system cancer; bladder cancer; pancreatic cancer;

XX KW cervical cancer; melanoma; leukaemia; hybridisation probe;

XX KW chromosome identification; chromosome mapping; gene mapping;

XX KW gene therapy; cytostatic; gene; ss.

XX OS Homo sapiens.

XX PN WO2004030615-A2.

XX PD 15-APR-2004.

XX PF 29-SEP-2003; 2003WO-US028547.

XX PR 02-OCT-2002; 2002US-0414971P.

XX PA (GETH) GENENTECH INC.

XX PI Wu TD, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

XX P-FSDB; ABM80804.

XX New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.

XX Claim 1; SEQ ID NO 2070; 7273pp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT nucleic acid of the invention

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 10.7 Length: 2053
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-12 (1-9) x ACN38510 (1-2053)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 Db 367 AACCTCTTCCTTACCGGCACACGCTG 393

RESULT 25

ADV35098

ID ADV35098 standard; cDNA; 2053 BP.

XX AC ADV35098;

XX AC ADV35098;

XX DT 10-FEB-2005 (first entry)

XX DE Human cDNA of an exemplary efficacy gene for BAD SeqID174.

XX KW human; ss: multi-parameter high throughput screening; MPHTS;

XX KW disease signature; neuropsychiatric; neurodegenerative; MPHTS;

XX KW bipolar affective disorder; BAD; autism; Parkinson's;

XX KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.

XX OS Homo sapiens.

XX FN US2003096264-A1.

XX PD 22-MAY-2003.

XX PF 18-JUN-2002; 2002US-00175523.

XX PR 18-JUN-2001; 2001US-0299151P.

XX PR 07-SEP-2001; 2001US-0317828P.

XX PR 25-SEP-2001; 2001US-0325150P.

XX PR 14-NOV-2001; 2001US-0333047P.

XX PR 18-JAN-2002; 2002US-0349936P.

XX PR 04-MAR-2002; 2002US-0361834P.

XX PA (PSYC-) PSYCHIATRIC GENOMICS INC.

XX PI Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;

XX PI Palfreyman M, Rajan P;

XX DR WPI; 2004-118903/12.

XX PT Identifying a compound that can treat disease or disorders, such as, a

XX PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises

XX PT determining the expression of one or more efficacy genes in a cell

XX PT contacted with the test compound.

XX PS Example 6; SEQ ID NO 174; 39pp; English.

XX CC This invention relates to a novel screening method identified as a multi-

XX CC parameter high throughput screening (MPHTS) assay. Specifically, it

XX CC refers to an assay that utilizes the disease signature of a plurality of

XX CC specific genes associated with a particular disease, and identifies

XX CC differential expression between those cells taken from individuals

XX CC affected by that disease and those that are not affected. The present

XX CC invention then describes the screening of candidate pharmaceutical

XX CC compounds to identify those that have a potential therapeutic benefit for

XX CC the treatment of neuropsychiatric and neurodegenerative disorders

XX CC including schizophrenia, bipolar affective disorder (BAD) and autism, as

XX CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds

XX CC of this invention exhibit various activities including neuroleptic,

XX CC nootropic, antimanic and antidepressant. Furthermore, the screening

XX CC method used in MPHTS will be automated, such that a large number of test

XX CC compounds may be rapidly screened with a minimal amount of labour and

XX CC effort. This polynucleotide is a human cDNA sequence of a gene that is

CC differentially expressed in the presence of a therapeutic compound and

CC represents an exemplary efficacy gene for bipolar affective disorder,

CC given in an exemplification of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 10.7 Length: 2053

Score: 46.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 13 Gaps: 0

US-10-774-176-12 (1-9) x ADV35098 (1-2053)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

Db 367 AACCTCTTCCTTACCGGCACACGCTG 393

RESULT 26

AAS87175

ID AAS87175 standard; cDNA; 2338 BP.

XX AC AAS87175;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #22979.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Dmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG22988.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity.

XX PS Claim 1; SEQ ID NO 22979; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX CC sequences. (I) is useful as hybridisation probes, polymerase chain

XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

XX CC and in recombinant production of (II). The polynucleotides are also used

XX CC in diagnostics as expressed sequence tags for identifying expressed

XX CC genes. (I) is useful in gene therapy techniques to restore normal

XX CC activity of (II) or to treat disease states involving (II). (II) is

XX CC useful for generating antibodies against it, detecting or quantitating a

XX CC polypeptide in tissue, as molecular weight markers and as a food

XX CC supplement. (II) and its binding partners are useful in medical imaging

XX CC of sites expressing (II). (I) and (II) are useful for treating disorders

XX CC involving aberrant protein expression or biological activity. The

XX CC polypeptide and polynucleotide sequences have applications in

XX CC diagnostics, forensics, gene mapping, identification of mutations

XX CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 12.5 Length: 2338
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0

US-10-774-176-12 (1-9) x AAS87175 (1-2338)

QY 1 AsnLeuPheLeuThrGlyAenGlnLeu 9
 DB 624 AACCTCTTCCTTACCGGCAACACGCTG 650

RESULT 27
 AAK94253
 ID AAK94253 standard; cDNA; 2359 BP.

XX AC AAK94253;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human full-length cDNA, SEQ ID NO: 2864.
 XX KW Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX OS Homo sapiens.

XX PN EP1130094-A2.
 XX PD 05-SEP-2001.
 XX PF 07-JUL-2000; 2000EP-00114089.
 XX PR 08-JUL-1999; 99JP-00194486.
 XX PR 11-JAN-2000; 2000JP-00118774.
 XX PR 02-MAY-2000; 2000JP-00183765.

XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 XX PI Wakamatsu A, Sugiyama T, Negai K, Kojima S, Otsuki T, Koga H;
 XX P-PSDB; AAM93333.

XX WPI; 2001-524255/58.
 XX DR P-PSDB; AAM93333.

XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
 XX in genetic manipulation.

XX PS Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.

XX CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO

SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 12.6 Length: 2359
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-12 (1-9) x AAK94253 (1-2359)

QY 1 AsnLeuPheLeuThrGlyAenGlnLeu 9
 DB 706 AACCTCTTCCTTACCGGCAACACGCTG 732

RESULT 28
 ADL30831
 ID ADL30831 standard; cDNA; 2359 BP.

XX AC ADL30831;
 XX DT 20-MAY-2004 (first entry)
 XX DE Full length human cDNA clone SeqID 2864.
 XX KW human; medicine; signal transduction; glycoprotein; transcription;
 XX KW oligo-capping method; ss; gene.
 XX OS Homo sapiens.

XX PN EP1396543-A2.
 XX PD 10-MAR-2004.
 XX PF 07-JUL-2000; 2003EP-00025638.

XX PR 08-JUL-1999; 99JP-00194486.
 XX PR 11-JAN-2000; 2000JP-00118774.
 XX PR 02-MAY-2000; 2000JP-00183865.
 XX PR 07-JUL-2000; 2000EP-00114089.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 XX PI Wakamatsu A, Sugiyama T, Negai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2004-204755/20.
 XX DR P-PSDB; ADL30832.

XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 XX length human cDNAs.

XX PS Example 1; SEQ ID NO 2864; 1340pp; English.

XX CC This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3',
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.

XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 12.6 Length: 2359
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0

DB: 12 Gaps: 0

US-10-774-176-12 (1-9) x ADL30831 (1-2359)

QY 1 AsnLeuPheLeuThrGlyAanGlnLeu 9
 |||||
 Db 706 AACCTTCTTCTTACCGGCAACGAGCTG 732

RESULT 29
 AAK94254
 ID AAK94254 standard; cDNA; 2361 BP.

XX AC AAK94254;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human full-length cDNA, SEQ ID NO: 2866.
 XX KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX OS Homo sapiens.
 XX PN EF1130094-A2.
 XX PD 05-SEP-2001.
 XX PF 07-JUL-2000; 2000EP-00114089.
 XX PR 08-JUL-1999; 99JP-00194486.
 XX PR 11-JAN-2000; 2000JP-00118774.
 XX PR 02-MAY-2000; 2000JP-00183765.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;
 XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX P-PSDB; AAM93334.
 XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
 XX PT in genetic manipulation.
 XX PS Claim 8; SEQ ID NO 2866; 1380pp + Sequence listing; English.
 XX CC The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesising the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 12.6 Length: 2361
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 Db: 4 Gaps: 0

US-10-774-176-12 (1-9) x AAK94254 (1-2361)

QY 1 AsnLeuPheLeuThrGlyAanGlnLeu 9
 |||||
 Db 708 AACCTTCTTCTTACCGGCAACGAGCTG 734

DB: 12 Gaps: 0

US-10-774-176-12 (1-9) x ADL30831 (1-2359)

QY 1 AsnLeuPheLeuThrGlyAanGlnLeu 9
 |||||
 Db 706 AACCTTCTTCTTACCGGCAACGAGCTG 732

RESULT 30
 ADI26162
 ID ADI26162 standard; cDNA; 2361 BP.

XX AC ADI26162;
 XX DT 22-APR-2004 (first entry)
 XX DE Human cDNA encoding protein that promotes STAT6 activation #64.
 XX KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX OS Homo sapiens.
 XX PN WO2003104277-A2.
 XX PD 18-DEC-2003.
 XX PF 05-JUN-2003; 2003WO-JP007123.
 XX PR 05-JUN-2002; 2002JP-00164257.
 XX PR 06-JUN-2002; 2002US-0385912P.
 XX PR 26-DEC-2002; 2002JP-00377326.
 XX PR 27-DEC-2002; 2002US-0436467P.
 XX PR 15-MAY-2003; 2003JP-00137505.
 XX PR 16-MAY-2003; 2003US-0470836P.
 XX PA (ASAH) ASAH KASEI KK.
 XX PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX WPI; 2004-122214/12.
 XX DR P-PSDB; ADI26163.
 XX PT New signal transducer and activator of transcription 6 activation
 XX PT promoting purified protein, for diagnosing and treating disease
 XX PT associated with activation/inhibition of transcription factor e.g.
 XX PT diabetes and cancer.
 XX PS Claim 4; SEQ ID NO 127; 1368pp; English.
 XX CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficiently promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:

ID ADI26160 standard; cDNA; 2557 BP.
 XX ADI26160;
 AC
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #63.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation; cancer;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 XX 05-JUN-2003; 2003WO-JP007123.
 XX
 XX 05-JUN-2002; 2002JP-00164257.
 PR
 PR 06-JUN-2002; 2002US-0385912P.
 PR
 PR 26-DEC-2002; 2002JP-00377326.
 PR
 PR 27-DEC-2002; 2002US-0436467P.
 PR
 PR 15-MAY-2003; 2003JP-00137505.
 PR
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 XX (ASAH) ASAH KASEI KK.
 XX
 PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX
 XX WPI; 2004-122214/12.
 DR
 DR P-PSDB; ADI26161.
 XX
 XX New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 XX Claim 4; SEQ ID NO 125; 1368pp; English.
 PS
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has effectively useful for screening
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 37.1 Length: 2557
 Score: 44.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1

Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.7% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-12 (1-9) x ADI26160 (1-2557)

QY 1 AsnLeuPheLeuThrGlyAsnGlnIleu 9
 DB 838 AACCTTTTCCTTACCGGCACACAGATG 864
 RESULT 34
 ADI26158
 ID ADI26158 standard; cDNA; 2557 BP.
 XX
 AC ADI26158;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #62.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 XX 05-JUN-2003; 2003WO-JP007123.
 XX
 XX 05-JUN-2002; 2002JP-00164257.
 PR
 PR 06-JUN-2002; 2002US-0385912P.
 PR
 PR 26-DEC-2002; 2002JP-00377326.
 PR
 PR 27-DEC-2002; 2002US-0436467P.
 PR
 PR 15-MAY-2003; 2003JP-00137505.
 PR
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 XX (ASAH) ASAH KASEI KK.
 XX
 PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX
 XX WPI; 2004-122214/12.
 DR
 DR P-PSDB; ADI26159.
 XX
 XX New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 XX Claim 4; SEQ ID NO 123; 1368pp; English.
 PS
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has effectively useful for screening
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX

CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.

XX
SQ Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores: 37.1 Length: 2557
Pred. No.: 44.00 Matches: 8
Score: 100.0% Conservative: 1
Percent Similarity: 100.0%
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.7% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-12 (1-9) x ADI26158 (1-2557)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 838 AACCTTTTCCTTACCGGCAACGATG 864

RESULT 35
ADO35939/c
ID ADO35939 standard; DNA; 2557 BP.

XX ADO35939;

XX 26-AUG-2004 (first entry)

XX Novel mouse gene sequence #612.

DE mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
KW viral disorder; ds; gene.

XX Mus sp.

XX WO2004046310-A2.

XX 03-JUN-2004.

XX 24-OCT-2003; 2003WO-US033948.

XX 15-NOV-2002; 2002US-0426916P.

XX 04-DEC-2002; 2002US-0431158P.

XX 05-DEC-2002; 2002US-0431445P.

XX 05-DEC-2002; 2002US-0431606P.

XX 09-JUN-2003; 2003US-0476621P.

XX 09-JUN-2003; 2003US-0476632P.

XX 08-JUL-2003; 2003US-0485217P.

XX 08-JUL-2003; 2003US-0485359P.

XX 08-AUG-2003; 2003US-0493332P.

XX 08-AUG-2003; 2003US-0493356P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee B, Hestir K, Hayaishizaki Y, Kamiya M;

XX WPI; 2004-431966/40.

XX New mouse nucleic acid molecules and polypeptides, useful for treating
PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
PT disease or thrombosis.

XX Claim 1; SEQ ID NO 612; 263pp; English.

XX The invention comprises 744 novel mouse DNA sequences (genes). The DNA
CC sequences of the invention are useful for treating cancer, psoriasis,
CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
CC immune disorders, bacterial disorders and viral disorders. The present
CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
CC present DNA sequence is not shown in the specification, but has been

CC retrieved from the WIPO website.

XX SQ Sequence 2557 BP; 610 A; 794 C; 688 G; 465 T; 0 U; 0 Other;

Alignment Scores: 37.1 Length: 2557
Pred. No.: 44.00 Matches: 8
Score: 100.0% Conservative: 1
Percent Similarity: 100.0%
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.7% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-12 (1-9) x ADO35939 (1-2557)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 63 AACCTTTTCCTTACCGGCAACGATG 37

RESULT 36
ABL27915
ID ABL27915 standard; DNA; 2820 BP.

XX ABL27915;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 35218.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li FWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 35218; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABBS72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2820 BP; 721 A; 684 C; 716 G; 699 T; 0 U; 0 Other;

Alignment Scores: 112 Length: 2820
Pred. No.: 42.00 Matches: 8
Score: 100.0% Conservative: 0
Percent Similarity: 100.0%
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0

DB: 4 Gaps: 0

US-10-774-176-12 (1-9) x ABL27915 (1-2820)

Qy 1 AsnLeuPheLeuThrGlyAsnGln 8
|||||

Db 864 AATCTATTCTCTCACTGGGACCAA 887
|||||

RESULT 37

ABL27914/c

ID ABL27914 standard; DNA; 5278 BP.

XX AC ABL27914;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 35215.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical; gene; ds.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX PS Claim 1; SEQ ID NO 35215; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 5278 BP; 1496 A; 1188 C; 1149 G; 1445 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	231	Length:	5278
Score:	42.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	91.3%	Indels:	0
DB:	4	Gaps:	0

US-10-774-176-12 (1-9) x ABL27914 (1-5278)

Qy 1 AsnLeuPheLeuThrGlyAsnGln 8
|||||

Db 3355 AATCTATTCTCTCACTGGGACCAA 3332
|||||

RESULT 38

ADA02798

IDA02798 standard; DNA; 52754 BP.

ADA02798;

06-NOV-2003 (first entry)

Human TNFSF11 carcinoma associated gene, SEQ ID NO:1316.

Human; carcinoma associated; oncogene; carcinoma; cancer; breast; prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening; gene; ds.

Homo sapiens.

WO2003057146-A2.

17-JUL-2003.

26-DEC-2002; 2002WO-US041414.

26-DEC-2001; 2001US-00035832.

(SAGR-) SAGRES DISCOVERY.

Morris DW;

WPI; 2003-587068/55.

New recombinant nucleic acid encoding carcinoma associated protein, useful for preparing compositions for treating carcinomas.

Claim 1; SEQ ID NO 1316; 245pp; English.

The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 52754 BP; 15093 A; 10533 C; 11422 G; 15706 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.49e+03	Length:	52754
Score:	41.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	88.9%	Mismatches:	0
Query Match:	89.1%	Indels:	0
DB:	9	Gaps:	0

US-10-774-176-12 (1-9) x ADA02798 (1-52754)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
:::|||||

Db 16134 CACCTATTCTTACAGGCATCAGCTA 16160
|||||

RESULT 39

ADB72536
ID ADB72536 standard; DNA; 52754 BP.
XX
AC ADB72536;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human TNFSF11 gene.
XX
KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Homo sapiens.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
WPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 364; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a human gene of the invention.
XX
SQ Sequence 52754 BP; 15093 A; 10533 C; 11422 G; 15706 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.49e+03 Length: 52754
Score: 41.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 89.1% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-12 (1-9) x ADB72536 (1-52754)
QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 16134 CACCTATTCTTACAGGCAATCAGCTA 16160
RESULT 40
ID ADC85278 standard; DNA; 52754 BP.
XX
AC ADC85278;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human Tnfsf11 genomic sequence.
XX
KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.

XX Homo sapiens.
XX WO2003045230-A2.
XX
PD 05-JUN-2003.
XX
PF 02-DEC-2002; 2002WO-US038582.
XX
PR 30-NOV-2001; 2001US-00997722.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
WPI; 2003-513603/48.
XX
PT New recombinant nucleic acid comprising a nucleotide sequence of any of
PT the carcinoma-associated (CA) genes, useful for screening for drug
PT candidates for diagnosing or treating carcinomas.
XX
PS Claim 1; SEQ ID NO 64; 983pp; English.
XX
CC The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC ADC85514 represent CA genes of the invention.
XX
SQ Sequence 52754 BP; 15093 A; 10533 C; 11422 G; 15706 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.49e+03 Length: 52754
Score: 41.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 89.1% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-12 (1-9) x ADC85278 (1-52754)
QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 16134 CACCTATTCTTACAGGCAATCAGCTA 16160
RESULT 41
ID ADM74393 standard; DNA; 52754 BP.
XX
AC ADM74393;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human carcinoma associated (CA) nucleic acid #31.
XX
KW Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
KW cytostatic.
XX
OS Homo sapiens.
XX
PN US2004072154-A1.
XX
PD 15-APR-2004.
XX
PF 30-NOV-2001; 2001US-00997722.
XX
PR 22-DEC-2000; 2000US-00747377.
PR 02-MAR-2001; 2001US-00798586.
XX
PA (MORR/) MORRIS D W.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
DR WPI; 2001-656860/75.
DR P-PSDB; ABB70141.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 37214; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 37091 BP; 10549 A; 7800 C; 7769 G; 10973 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 5.98e+03 Length: 37091
Score: 40.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 87.0% Indels: 0
DB: 4 Gaps: 0
US-10-774-176-12 (1-9) x ABL14244 (1-37091)
QY 1 AsnLeuPheLeuThrGlyAenGlnLeu 9
DB 78 AACATTTTTCGAGTGGGAATCACTT 104
RESULT 44
ADP77343_11
Continuation (12 of 20) of ADP77343 from base 1100001 (Lactic acid bacteria Lactobacillus
WP Sequence split into 20 fragments LOCUS ADP77343 Accession Adp77343
WP Fragment Name Begin End
WP ADP77343_00 1 110000
WP ADP77343_01 100001 210000
WP ADP77343_02 200001 310000
WP ADP77343_03 300001 410000
WP ADP77343_04 400001 510000
WP ADP77343_05 500001 610000
WP ADP77343_06 600001 710000
WP ADP77343_07 700001 810000
WP ADP77343_08 800001 910000
WP ADP77343_09 900001 1010000
WP ADP77343_10 1000001 1110000
WP ADP77343_11 1100001 1210000
WP ADP77343_12 1200001 1310000

WP ADP77343_13 1300001 1410000
WP ADP77343_14 1400001 1510000
WP ADP77343_15 1500001 1610000
WP ADP77343_16 1600001 1710000
WP ADP77343_17 1700001 1810000
WP ADP77343_18 1800001 1910000
WP ADP77343_19 1900001 1983043
Alignment Scores:
Pred. No.: 2.11e+04 Length: 110000
Score: 40.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 87.0% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-12 (1-9) x ADP77343_11 (1-110000)
QY 1 AsnLeuPheLeuThrGlyAenGln 8
DB 76630 AATATATTCTCACTGGGAATCAG 76653
RESULT 45
ABQ42606/C
ID ABQ42606 standard; DNA; 792 BP.
XX
XX AC ABQ42606;
XX
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29197.
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200218632-A2.
XX
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP010074.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PR 05-SEP-2000; 2000DE-01044543.
XX
XX PA (EPIG-) EPIGENOMICS AG.
XX
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
XX DR WPI; 2002-371829/40.
XX
XX PT Determining the degree of cytosine methylation in genomic DNA, useful for
XX diagnosis and prognosis, comprises selective hybridization of amplicons
XX from chemically treated DNA.
XX
XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX CC This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one member,
XX of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
XX degree of hybridisation to both classes is determined from the label on
XX the amplicon. From the ratio of labels hybridised to the two classes of
XX oligomers, the degree of methylation is calculated. The method is used:
XX (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
XX and of a wide range of diseases, e.g. cancer, disorders of the central
XX nervous, cardiovascular, gastrointestinal and respiratory systems etc.,

CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

XX SQ Sequence 792 BP; 119 A; 94 C; 297 G; 282 T; 0 U; 0 Other;

Alignment Scores: 112 Length: 792
 Pred. No.: 39.00 Matches: 8
 Score: 39.00 Conservatives: 0
 Percent Similarity: 88.9% Mismatches: 1
 Best Local Similarity: 88.9% Indels: 0
 Query Match: 84.8% Gaps: 0
 DB: 6

US-10-774-176-12 (1-9) x ABQ42606 (1-792)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

DB 166 AACCTCTCTCTTACCGACACCAACTA 140

RESULT 46

ABQ42607

ID ABQ42607 standard; DNA; 792 BP.

XX AC ABQ42607;

DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29198.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX DR Determining the degree of cytosine methylation in genomic DNA, useful for
 XX diagnosis and prognosis, comprises selective hybridization of amplicons
 XX from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs

CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

XX SQ Sequence 792 BP; 282 A; 297 C; 94 G; 119 T; 0 U; 0 Other;

Alignment Scores: 112 Length: 792
 Pred. No.: 39.00 Matches: 8
 Score: 39.00 Conservatives: 0
 Percent Similarity: 88.9% Mismatches: 1
 Best Local Similarity: 88.9% Indels: 0
 Query Match: 84.8% Gaps: 0
 DB: 6

US-10-774-176-12 (1-9) x ABQ42607 (1-792)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

DB 627 AACCTCTCTCTTACCGACACCAACTA 653

RESULT 47

ABQ42584/C

ID ABQ42584 standard; DNA; 795 BP.

XX AC ABQ42584;

DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29175.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX DR Determining the degree of cytosine methylation in genomic DNA, useful for
 XX diagnosis and prognosis, comprises selective hybridization of amplicons
 XX from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of

CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention
 XX
 SQ Sequence 795 BP; 132 A; 94 C; 313 G; 256 T; 0 U; 0 Other;

Alignment Scores: 113 Length: 795
 Pred. No.: 39.00 Matches: 8
 Score: 88.9% Conservative: 0
 Percent Similarity: 88.9% Mismatches: 1
 Best Local Similarity: 84.8% Indels: 0
 Query Match: 6 Gaps: 0
 DB: 6

US-10-774-176-12 (1-9) x ABQ42584 (1-795)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 DB 256 AACCTCTCTCTTACCGACACCAACTA 230
 |||||

RESULT 49

ABQ42585
 ID ABQ42585 standard; DNA; 795 BP.

XX AC ABQ42585;

DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29176.
 KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX WO200218632-A2.

XX PD 07-MAR-2002.

XX PP 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.

XX PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX CC This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the

CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention
 XX

SQ Sequence 795 BP; 256 A; 313 C; 94 G; 132 T; 0 U; 0 Other;

Alignment Scores: 113 Length: 795
 Pred. No.: 39.00 Matches: 8
 Score: 88.9% Conservative: 0
 Percent Similarity: 88.9% Mismatches: 1
 Best Local Similarity: 84.8% Indels: 0
 Query Match: 6 Gaps: 0
 DB: 6

US-10-774-176-12 (1-9) x ABQ42585 (1-795)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 DB 540 AACCTCTCTCTTACCGACACCAACTA 566
 |||||

RESULT 49

ADS48405

ID ADS48405 standard; cDNA; 1874 BP.

XX AC ADS48405;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polynucleotide #3148.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.

XX OS Bacteria.

XX PN US20032333675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 26835; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 1874 BP; 446 A; 503 C; 547 G; 378 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 500 Length: 1874
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 82.6% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-12 (1-9) x ADS48405 (1-1874)

Oy 1 AsnLeuPheLeuThrGlyAenGlnLeu 9
||| :|||
Db 100 AACACTTATCTGACAGGTATACGCTC 126

RESULT 50

AAK70925

ID AAK70925 standard; DNA; 11646 BP.

AC AAK70925;

DT 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25737.

DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; Gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUL-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-02411787P.
PR 20-OCT-2000; 2000US-02411808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 25737; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to

CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
XX
SQ Sequence 11646 BP; 2490 A; 3259 C; 3106 G; 2791 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.18e+03	Length:	11646
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	87.5%	Mismatches:	0
Query Match:	82.6%	Indels:	0
DB:	4	Gaps:	0

US-10-774-176-12 (1-9) x AAK70925 (1-11646)

Qy 1 AsnLeuPheLeuThrGlyAsnGln 8

Db 7229 AATTGTTCCTCTCTGGGAATCAA 7252

Search completed: April 25, 2006, 12:37:47
Job time : 328.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n.model

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignments)
171.290 Million cell updates/sec

Title: US-10-774-176-12

Perfect score: 46

Sequence: 1 NLFITGNQL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5983141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10774176/runat_24042006_165114_19197/app.query.fasta_1
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04
-USER=US10774176 @CGN 1.1 6765 @runat_24042006_165114_19197 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARM TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_ets:*
11: gb_ey:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	927	6	AX829164 Sequence
2	46	100.0	1260	6	AX467373 Sequence
3	46	100.0	1260	6	AX821533 Sequence

4	46	100.0	1260	6	AX821548
5	46	100.0	1263	6	BD249731
6	46	100.0	1263	6	AX025011
7	46	100.0	1263	6	AX149553
8	46	100.0	1263	6	AX316086
9	46	100.0	1263	6	AX467371
10	46	100.0	2053	6	CQ731678
11	46	100.0	2053	8	H55740A
12	46	100.0	2359	6	BD127282
13	46	100.0	2359	6	CQ782724
14	46	100.0	2359	8	AK074786
15	46	100.0	2361	6	BD127283
16	46	100.0	2361	6	CQ782726
17	46	100.0	2361	6	AX961916
18	46	100.0	2361	8	AK074790
19	46	100.0	2379	8	BC037161
20	46	100.0	2714	8	AB168308
21	46	100.0	5551	8	HSA012159
22	46	100.0	110000	15	AP008208_073
23	46	100.0	121909	8	HSJ492P14
24	46	100.0	136267	15	AP005756
25	44	95.7	1281	6	BD249732
26	44	95.7	1281	6	AX025012
27	44	95.7	1281	6	AX316087
28	44	95.7	2333	9	AF063939
29	44	95.7	2361	9	BC087011
30	44	95.7	2423	9	BC058198
31	44	95.7	2557	6	AX961912
32	44	95.7	2557	6	AX961914
33	44	95.7	7942	9	MMU012160
34	44	95.7	85624	8	AL139393
35	44	95.7	167046	9	AC158516
36	44	95.7	187098	14	AC015984
37	44	95.7	210237	14	AC128294
38	44	95.7	239076	14	AC106962
39	42	91.3	1673	2	AY071256
40	42	91.3	2820	6	CQ611355
41	42	91.3	5278	6	CQ611354
42	42	91.3	68662	14	AC018828
43	42	91.3	77707	14	AC014787
44	42	91.3	121242	9	AC123802
45	42	91.3	129302	14	AC083918
46	42	91.3	144356	8	AC021486
47	42	91.3	161674	14	AC015567
48	42	91.3	165591	8	AC019239
49	42	91.3	175819	9	AC102290
50	42	91.3	177480	2	AC008287
51	42	91.3	179908	9	AC125344
52	42	91.3	225974	2	AB003778
53	42	91.3	255004	14	AC112871
54	42	91.3	258015	14	AC131816
55	41	89.1	52754	6	AX695689
56	41	89.1	164587	14	AC023001
57	41	89.1	166765	8	AC112496
58	41	89.1	171727	8	HSB138P15
59	41	89.1	178918	9	AC120159
60	41	89.1	200724	8	AL139382
61	40	87.0	698	10	BV625221
62	40	87.0	794	15	AY233204
63	40	87.0	1611	2	AF303112
64	40	87.0	37091	6	CQ590849
65	40	87.0	52726	14	AC017352
66	40	87.0	86908	5	CR769770
67	40	87.0	88095	15	ATP13G24
68	40	87.0	104339	14	AF235105
69	40	87.0	104967	14	CR955010
70	40	87.0	105335	8	AC107374
71	40	87.0	149866	8	AC025627
72	40	87.0	156168	8	AC124917
73	40	87.0	158456	8	AC092198
74	40	87.0	162587	8	AC092805
75	40	87.0	173751	14	AL590292
76	40	87.0	174208	14	AC015935

AX821548 Sequence
BD249731 Polypepti
AX025011 Sequence
AX149553 Sequence
AX316086 Sequence
AX467371 Sequence
CQ731678 Sequence
Z29083 Homo sapien
BD127282 Primer fo
CQ782724 Sequence
AK074786 Homo sapi
BD127283 Primer fo
CQ782726 Sequence
AX961916 Sequence
AK074790 Homo sapi
BC037161 Homo sapi
AB168308 Macaca fa
AJ012159 Homo sapi
Continuation (74 o
AL121977 Human DNA
AP005756 Oryza sat
BD249732 Polypepti
AX025012 Sequence
AX316087 Sequence
AF063939 Rattus no
BC087011 Rattus no
BC058198 Mus muscu
AX961912 Sequence
AX961914 Sequence
AJ012160 Mus muscu
AL139393 Human DNA
AC158516 Mus muscu
AC015984 Homo sapi
AC128294 Rattus no
AC106962 Rattus no
AY071256 Drosophil
CQ611355 Sequence
CQ611354 Sequence
AC018828 Mus muscu
AC014787 Drosophil
AC123802 Mus muscu
AC083918 Homo sapi
AC021486 Homo sapi
AC015567 Homo sapi
AC019239 Homo sapi
AC102290 Mus muscu
AC008287 Drosophil
AC125344 Mus muscu
AB003778 Drosophil
AC112871 Rattus no
AC131816 Rattus no
AX695689 Sequence
AC023001 Homo sapi
AC112496 Homo sapi
AL121869 Human DNA
AC120159 Mus muscu
AL139382 Human DNA
BV625221 S216P6046
AY233204 Vitia vin
AF303112 Physarum
CQ590849 Sequence
AC017352 Drosophil
CR769770 Zebrafish
AL133421 Arabidops
AF235105 Homo sapi
CR955010 Medicago
AC107374 Homo sapi
AC025627 Homo sapi
AC124917 Homo sapi
AC092198 Homo sapi
AC092805 Homo sapi
AL590292 Homo sapi
AC015935 Homo sapi

C 77	40	87.0	188765	8	AC118755	AC118755 Homo sapi	150	38	82.6	142924	14	CR762428	CR762428 Danio rer
C 78	40	87.0	189557	2	AC007356	AC007356 Drosophila	151	38	82.6	146327	8	AL162253	AL162253 Human DNA
C 79	40	87.0	192137	8	AC087501	AC087501 Homo sapi	152	38	82.6	147088	9	AC102443	AC102443 Mus muscu
C 80	40	87.0	195900	14	AC021164	AC021164 Homo sapi	c 153	38	82.6	150242	8	AC006195	AC006195 Homo sapi
C 81	40	87.0	228883	9	AC127173	AC127173 Mus muscu	154	38	82.6	150242	6	CS086351	CS086351 Sequence
C 82	40	87.0	230689	2	AC159326	AC159326 Mus muscu	155	38	82.6	152222	9	AC158544	AC158544 Mus muscu
C 83	40	87.0	234168	2	AE003821	AE003821 Drosophila	156	38	82.6	153852	14	AC009694	AC009694 Homo sapi
C 84	40	87.0	305128	1	AE017203	AE017203 Lactobaci	157	38	82.6	155268	9	AC132332	AC132332 Mus muscu
C 85	40	87.0	349980	6	AX926714	AX926714 Sequence	158	38	82.6	156128	5	BX324195	BX324195 Zebrafish
C 86	39	84.8	293	9	AF114822	AF114822 Clethrion	c 159	38	82.6	157119	5	EX276185	EX276185 Zebrafish
C 87	39	84.8	593	8	AF463548	AF463548 Gorilla g	160	38	82.6	157159	14	AC026279	AC026279 Homo sapi
C 88	39	84.8	88236	14	AP007807	AP007807 Lotus cor	161	38	82.6	157590	8	AC087525	AC087525 Homo sapi
C 89	39	84.8	88345	14	CR854913_3	Continuation (4 of	162	38	82.6	152202	14	AC158340	AC158340 Mus muscu
C 90	39	84.8	91071	14	AC022589	AC022589 Homo sapi	163	38	82.6	162784	15	AP006860	AP006860 Oryza sat
C 91	39	84.8	105600	14	AP007364	AP007364 Lotus cor	164	38	82.6	162784	8	AC083862	AC083862 Homo sapi
C 92	39	84.8	110000	14	AC111775_1	Continuation (2 of	165	38	82.6	165004	14	AC060826	AC060826 Homo sapi
C 93	39	84.8	110000	14	AC111775_1	Continuation (10 o	166	38	82.6	165990	14	AC093287	AC093287 Homo sapi
C 94	39	84.8	110000	14	AC156264_09	Continuation (10 o	167	38	82.6	169199	8	AC090497	AC090497 Homo sapi
C 95	39	84.8	130117	8	AC004907	AC004907 Homo sapi	168	38	82.6	169404	8	AL353751	AL353751 Human DNA
C 96	39	84.8	139173	8	AC009140	AC009140 Homo sapi	169	38	82.6	169404	14	AC149250	AC149250 Papio anu
C 97	39	84.8	157308	14	AC079351	AC079351 Homo sapi	170	38	82.6	171024	14	AC149102	AC149102 Papio anu
C 98	39	84.8	167776	14	AC092294	AC092294 Homo sapi	c 171	38	82.6	171039	14	AC090312	AC090312 Homo sapi
C 99	39	84.8	168032	14	AC022785	AC022785 Homo sapi	172	38	82.6	171438	8	AC090312	AC090312 Homo sapi
C 100	39	84.8	168825	14	CR854985	CR854985 Danio rer	c 173	38	82.6	171456	8	AC023591	AC023591 Homo sapi
C 101	39	84.8	170655	14	AP001012	AP001012 Homo sapi	c 174	38	82.6	172413	14	AP001449	AP001449 Homo sapi
C 102	39	84.8	175794	8	AC147031	AC147031 Pan trogl	c 175	38	82.6	172413	15	AC027659	AC027659 Oryza sat
C 103	39	84.8	176340	9	AC116892	AC116892 Mus muscu	176	38	82.6	174291	14	AC118568	AC118568 Gallus ga
C 104	39	84.8	178029	8	AC106037	AC106037 Homo sapi	177	38	82.6	175908	14	AP001812	AP001812 Homo sapi
C 105	39	84.8	183002	8	AC138304	AC138304 Homo sapi	178	38	82.6	177744	14	AC073954	AC073954 Homo sapi
C 106	39	84.8	183098	14	CR339057	CR339057 Danio rer	179	38	82.6	180100	14	AC150382	AC150382 Callicebu
C 107	39	84.8	183986	14	AC122064	AC122064 Rattus no	c 180	38	82.6	180423	9	AC125355	AC125355 Mus muscu
C 108	39	84.8	187316	8	AL139340	AL139340 Homo sapi	c 181	38	82.6	180922	5	AL627168	AL627168 Zebrafish
C 109	39	84.8	194874	14	AC080090	AC080090 Homo sapi	c 182	38	82.6	180922	14	AC069269	AC069269 Homo sapi
C 110	39	84.8	200965	14	AC020748	AC020748 Homo sapi	c 183	38	82.6	180999	14	AC150610	AC150610 Callithri
C 111	39	84.8	202421	8	AC091038	AC091038 Homo sapi	c 184	38	82.6	184340	14	AC155201	AC155201 Callithri
C 112	39	84.8	203628	9	AC160633	AC160633 Mus muscu	c 185	38	82.6	185550	14	CR354564	CR354564 Danio rer
C 113	39	84.8	231655	14	AC111681	AC111681 Rattus no	c 186	38	82.6	185550	14	AC150382	AC150382 Rattus no
C 114	39	84.8	232458	14	AC103209	AC103209 Rattus no	c 187	38	82.6	189229	9	AC157656	AC157656 Mus muscu
C 115	39	84.8	237279	14	AC128607	AC128607 Rattus no	c 188	38	82.6	193635	8	AC099557	AC099557 Homo sapi
C 116	39	84.8	244377	14	AC110661	AC110661 Rattus no	c 189	38	82.6	196501	8	AC005908	AC005908 Homo sapi
C 117	39	84.8	245450	14	AC125586	AC125586 Rattus no	c 190	38	82.6	198393	9	AL773531	AL773531 Mouse DNA
C 118	39	84.8	249450	14	AC108355	AC108355 Rattus no	c 191	38	82.6	202065	14	AC146397	AC146397 Pan trogl
C 119	39	84.8	256046	14	AC096263	AC096263 Rattus no	c 192	38	82.6	203315	14	AC149169	AC149169 Papio anu
C 120	39	84.8	338015	14	AC131401	AC131401 Rattus no	c 193	38	82.6	203518	14	AC123824	AC123824 Mus muscu
C 121	38	82.6	1207	15	AV138789	AV138789 Rhizopus	c 194	38	82.6	205607	9	AC123824	AC123824 Mus muscu
C 122	38	82.6	1981	15	STFELCIGEN	X93564 S. tuberosum	c 195	38	82.6	206029	8	AC146144	AC146144 Pan trogl
C 123	38	82.6	2183	15	DQ113467	DQ113467 Neurospor	c 196	38	82.6	206050	14	AC110646	AC110646 Rattus no
C 124	38	82.6	2195	15	DQ113456	DQ113456 Neurospor	c 197	38	82.6	207531	14	AC133061	AC133061 Rattus no
C 125	38	82.6	2195	15	DQ113457	DQ113457 Neurospor	c 198	38	82.6	211770	14	AC159578	AC159578 Papio anu
C 126	38	82.6	2195	15	DQ113458	DQ113458 Neurospor	c 199	38	82.6	218909	14	AC018806	AC018806 Homo sapi
C 127	38	82.6	3891	15	AY692025	AY692025 Cryphonet	c 200	38	82.6	226040	14	AC110162	AC110162 Mus muscu
C 128	38	82.6	7800	13	AY753327	AY753327 Shrimp wh	c 201	38	82.6	228414	14	AC128563	AC128563 Rattus no
C 129	38	82.6	32958	14	AC110295	AC110295 Homo sapi	c 202	38	82.6	228453	14	AC163553	AC163553 Bos tauru
C 130	38	82.6	65223	14	AC104242	AC104242 Homo sapi	c 203	38	82.6	239747	14	AC120706	AC120706 Rattus no
C 131	38	82.6	78095	8	AP000759	AP000759 Homo sapi	c 204	38	82.6	240339	14	AC109949	AC109949 Rattus no
C 132	38	82.6	79684	6	AX695965	AX695965 Sequence	c 205	38	82.6	242813	14	AC095386	AC095386 Rattus no
C 133	38	82.6	80117	8	AC055813	AC055813 Homo sapi	c 206	38	82.6	253402	8	AC008534	AC008534 Homo sapi
C 134	38	82.6	84129	15	AB013392	AB013392 Arabidops	c 207	38	82.6	255964	14	AC120297	AC120297 Rattus no
C 135	38	82.6	85189	14	AC166588	AC166588 Bos tauru	c 208	38	82.6	262962	14	AC105538	AC105538 Rattus no
C 136	38	82.6	95356	8	AL139112	AL139112 Human DNA	c 209	38	82.6	273712	14	AC096108	AC096108 Rattus no
C 137	38	82.6	98985	8	AC110809	AC110809 Homo sapi	c 210	38	82.6	280245	14	AC103319	AC103319 Rattus no
C 138	38	82.6	106542	14	AC153491	AC153491 Mus muscu	c 211	38	82.6	305192	13	AE017093	AE017093 Oryza sat
C 139	38	82.6	110000	2	CP000081_00	CP000081 Leishmani	c 212	38	82.6	307287	15	AF404570	AF404570 Shrimp wh
C 140	38	82.6	110000	4	AC132794_3	Continuation (4 of	c 213	38	82.6	315761	14	AL158079	AL158079 Homo sapi
C 141	38	82.6	110000	15	AP008216_136	Continuation (137	c 214	37	80.4	240	6	AR269554	AR269554 Sequence
C 142	38	82.6	110000	15	AP008216_137	Continuation (137	c 215	37	80.4	278	6	AR247061	AR247061 Sequence
C 143	38	82.6	110000	15	AP008216_138	Continuation (137	c 216	37	80.4	332	10	AB140600	AB140600 Homo sapi
C 144	38	82.6	110000	15	AP008216_139	Continuation (137	c 217	37	80.4	390	6	AX309228	AX309228 Sequence
C 145	38	82.6	110000	15	AP008216_140	Continuation (105	c 218	37	80.4	612	10	BV375146	BV375146 S231P6294
C 146	38	82.6	122814	14	AC136505	AC136505 Medicago	c 219	37	80.4	675	6	CQ733561	CQ733561 Sequence
C 147	38	82.6	126887	8	AL353146	AL353146 Human DNA	c 220	37	80.4	760	6	AR494810	AR494810 Sequence
C 148	38	82.6	131457	8	AL356600	AL356600 Human DNA	c 221	37	80.4	760	6	AX281095	AX281095 Sequence
C 149	38	82.6	139076	8	AC004147	AC004147 Homo sapi	c 222	37	80.4	769	1	AF496093	AF496093 Lactobaci

223	37	80.4	930	6	AR547908	Sequence	AR547908	Sequence	296	37	80.4	80968	5	EX927386
224	37	80.4	991	6	AR227460	Sequence	AR227460	Sequence	c 297	37	80.4	81672	15	AB020755
225	37	80.4	1059	6	BD227283	Secreted	BD227283	Secreted	c 298	37	80.4	81891	8	AC108076
c 226	37	80.4	1507	1	BAC506092	Sequence	D28169	Bacillus th	c 299	37	80.4	82316	15	ATT22K7
	37	80.4	1680	6	AX506092	Sequence	AX506092	Sequence	300	37	80.4	84832	14	AC161673
228	37	80.4	1686	6	EL1091	cDNA coding	EL1091	cDNA coding	301	37	80.4	89290	8	CNS01DVY
229	37	80.4	1713	5	BC099984	Danio rer	BC099984	Danio rer	302	37	80.4	89733	14	AC165028
230	37	80.4	1746	6	AX815466	Sequence	AX815466	Sequence	303	37	80.4	90858	14	AP008185
231	37	80.4	1746	15	AY040054	Brassicica	AY040054	Brassicica	304	37	80.4	95111	15	AYF27K19
232	37	80.4	1746	15	AF108123	Arabidops	AF108123	Arabidops	c 305	37	80.4	97500	4	AT227782
233	37	80.4	1765	5	AB097825	Danio rer	AB097825	Danio rer	306	37	80.4	97665	14	AP007411
234	37	80.4	1776	15	AY053422	Arabidops	AY053422	Arabidops	c 307	37	80.4	98441	15	AP007252
235	37	80.4	1794	15	BT008358	Arabidops	BT008358	Arabidops	308	37	80.4	100886	14	AP003831
c 236	37	80.4	1828	6	CQ842192	Sequence	CQ842192	Sequence	c 309	37	80.4	106700	8	AY451126
	37	80.4	1828	8	AK125218	Homo sapi	AK125218	Homo sapi	c 310	37	80.4	107098	15	AP006660
238	37	80.4	1894	15	GMU41473	Glycine max	U41473	Glycine max	c 311	37	80.4	110000	1	AE017220
239	37	80.4	1940	15	STPLC2	S.tuberosum	X94183	S.tuberosum	312	37	80.4	110000	1	AE017225
240	37	80.4	1946	15	AF434168	Arabidops	AF434168	Arabidops	313	37	80.4	110000	1	AE017334
241	37	80.4	1961	15	ATHATPLC1	Arabidops	D38544	Arabidops	314	37	80.4	110000	1	AE017355
242	37	80.4	1980	15	AY093217	Arabidops	AY093217	Arabidops	315	37	80.4	110000	1	CP000001
243	37	80.4	2006	15	VUUR5250	Vigna ungu	U85250	Vigna ungu	c 316	37	80.4	110000	14	AC110642
244	37	80.4	2009	15	STPLC3	S.tuberosum	X94289	S.tuberosum	c 317	37	80.4	110000	14	AC110642
245	37	80.4	2013	15	NRV11931	N.rustica m	Y11931	N.rustica m	c 318	37	80.4	110000	14	AC127805
246	37	80.4	2032	15	AF360206	Arabidops	AF360206	Arabidops	319	37	80.4	110000	14	AC359456
247	37	80.4	2038	15	DSA291467	Digitaria	AJ291467	Digitaria	c 320	37	80.4	110000	15	AP008215
248	37	80.4	2066	15	NRPHOSLPC	N.rustica m	X95877	N.rustica m	c 321	37	80.4	110000	15	AP008216
249	37	80.4	2079	15	GRW25027	Glycine max	U25027	Glycine max	322	37	80.4	110000	15	AP008218
250	37	80.4	2098	15	AF223351	Nicotiana	AF223351	Nicotiana	c 323	37	80.4	110000	15	AP008218
251	37	80.4	2114	15	FSPLC	Fisum sativ	Y15233	Fisum sativ	c 324	37	80.4	110000	15	AP008218
252	37	80.4	2134	15	GMU41474	Glycine max	U41474	Glycine max	325	37	80.4	110000	15	AP008218
253	37	80.4	2139	6	AR494816	Sequence	AR494816	Sequence	c 326	37	80.4	110000	15	AC145127
254	37	80.4	2139	6	AX281100	Sequence	AX281100	Sequence	327	37	80.4	110000	15	AP008208
255	37	80.4	2164	15	ATHATPLC2	Arabidops	D50804	Arabidops	328	37	80.4	110000	15	AP008209
256	37	80.4	2170	15	GMU41475	Glycine max	U41475	Glycine max	329	37	80.4	110000	15	AP008211
257	37	80.4	2207	15	AF332874	Oryza sat	AF332874	Oryza sat	c 330	37	80.4	110000	15	AP008213
258	37	80.4	2218	15	AK064924	Sequence	AK064924	Oryza sat	c 331	37	80.4	111350	8	ALI57814
259	37	80.4	2276	6	AR494815	Sequence	AR494815	Sequence	c 332	37	80.4	111402	9	AP003154
260	37	80.4	2276	6	AX281100	Sequence	AX281100	Sequence	c 333	37	80.4	115479	4	AC144643
261	37	80.4	2307	15	AK070452	Oryza sat	AK070452	Oryza sat	c 334	37	80.4	116177	8	HS265A22
262	37	80.4	2323	15	AY150803	Arabidops	AY150803	Arabidops	c 335	37	80.4	118018	15	AP005675
263	37	80.4	2374	15	AK119748	Oryza sat	AK119748	Oryza sat	c 336	37	80.4	123576	8	AC015969
264	37	80.4	2394	15	AF223573	Nicotiana	AF223573	Nicotiana	c 337	37	80.4	123708	8	HS190A9
c 265	37	80.4	2397	5	BC075234	Xenopus l	BC075234	Xenopus l	c 338	37	80.4	127426	15	AC137075
	37	80.4	2420	6	AK835218	Sequence	AK835218	Sequence	c 339	37	80.4	127426	15	CNS08CAB
267	37	80.4	2420	8	AK098148	Homo sapi	AK098148	Homo sapi	340	37	80.4	128223	15	AP005643
268	37	80.4	2430	15	AB114834	Physcomit	AB114834	Physcomit	c 341	37	80.4	128726	8	ALI36458
269	37	80.4	2712	15	AY394079	Vigna rad	AY394079	Vigna rad	c 342	37	80.4	130195	14	AC090170
270	37	80.4	2860	15	AK120189	Oryza sat	AK120189	Oryza sat	c 343	37	80.4	131217	13	AY528864
271	37	80.4	2916	15	AK072141	Oryza sat	AK072141	Oryza sat	c 344	37	80.4	131441	15	CNS08CDA
272	37	80.4	2965	15	AF280748	Pisum sat	AF280748	Pisum sat	c 345	37	80.4	135323	15	CNS08CBQ
273	37	80.4	3009	8	AK026747	Homo sapi	AK026747	Homo sapi	346	37	80.4	135401	9	AL772324
274	37	80.4	3010	6	AX780392	Sequence	AX780392	Sequence	c 347	37	80.4	137559	14	CR936948
c 275	37	80.4	3252	1	BAJ10138	Bacillus	AJ010138	Bacillus	c 348	37	80.4	138368	15	CNS08CB4
	37	80.4	3418	6	CQ491502	Sequence	CQ491502	Sequence	349	37	80.4	138461	14	AC113576
c 277	37	80.4	3418	6	CQ497386	Sequence	CQ497386	Sequence	c 350	37	80.4	138594	14	AC160601
	37	80.4	3471	15	ATU76423	Arabidops	U76423	Arabidops	351	37	80.4	140041	15	CR382279
278	37	80.4	5213	15	AY394078	Vigna rad	AY394078	Vigna rad	c 352	37	80.4	140462	15	AC145219
280	37	80.4	6443	15	AY059631	Medicago s	AY059631	Medicago s	c 353	37	80.4	141751	8	BS000019
281	37	80.4	7555	2	AB217914	Lymnaea s	AB217914	Lymnaea s	354	37	80.4	141949	9	AC111039
282	37	80.4	12327	15	AC007658	Arabidops	AC007658	Arabidops	c 355	37	80.4	143436	8	AC004456
c 283	37	80.4	20653	1	AE008836	Salmonell	AE008836	Salmonell	c 356	37	80.4	144455	15	AC129720
	37	80.4	29605	15	AC074395	Arabidops	AC074395	Arabidops	c 357	37	80.4	146002	8	CNS01RGD
285	37	80.4	35653	8	AL445644	Human DNA	AL445644	Human DNA	c 358	37	80.4	147728	15	AP005644
286	37	80.4	37957	2	CRC04C11	Caenorhabdi	Z72501	Caenorhabdi	c 359	37	80.4	149022	14	AC135182
c 287	37	80.4	39989	14	AC091136	Homo sapi	AC091136	Homo sapi	360	37	80.4	151544	14	AC1321650
	37	80.4	46626	15	AC007293	Arabidops	AC007293	Arabidops	361	37	80.4	151841	5	EX908391
288	37	80.4	49839	14	AC166963	Bos tauru	AC166963	Bos tauru	c 362	37	80.4	152061	14	AC087849
289	37	80.4	57991	15	AC018721	Arabidops	AC018721	Arabidops	c 363	37	80.4	153642	14	AC160889
290	37	80.4	5946	14	AC012427	Homo sapi	AC012427	Homo sapi	c 364	37	80.4	153919	14	AL137864
291	37	80.4	65912	14	AC027511	Homo sapi	AC027511	Homo sapi	c 365	37	80.4	154713	8	AC133361
292	37	80.4	68990	14	AC087662	Homo sapi	AC087662	Homo sapi	c 366	37	80.4	155332	8	AC109581
293	37	80.4	69608	14	AC036201	Homo sapi	AC036201	Homo sapi	c 367	37	80.4	155517	9	AC158149
294	37	80.4	76202	14	AP008119	Lotus cor	AP008119	Lotus cor	c 368	37	80.4	155985	8	AC124467
295	37	80.4												

C 369	37	80.4	156415	14	AC157871	AC157871 Loxodonta	C 442	37	80.4	221913	9	AC148972	AC148972 Mus muscu
C 370	37	80.4	156654	15	AC074283	AC074283 Oryza sat	C 443	37	80.4	222741	9	AC140985	AC140985 Mus muscu
C 371	37	80.4	158180	14	AC100781	AC100781 Homo sapi	C 444	37	80.4	223449	14	AC112864	AC112864 Rattus no
C 372	37	80.4	158884	5	BX248399	BX248399 Zebrafish	C 445	37	80.4	223901	14	AC152811	AC152811 Bos tauru
C 373	37	80.4	158960	8	AL354835	AL354835 Human DNA	C 446	37	80.4	223954	14	AC103280	AC103280 Rattus no
C 374	37	80.4	161297	8	AC146076	AC146076 Pan trogl	C 447	37	80.4	225388	14	AC162984	AC162984 Bos tauru
C 375	37	80.4	161343	14	AC144869	AC144869 Pan trogl	C 448	37	80.4	225388	14	AC151061	AC151061 Bos tauru
C 376	37	80.4	162021	14	AC152138	AC152138 Daeypus n	C 449	37	80.4	226806	14	AC106207	AC106207 Rattus no
C 377	37	80.4	163285	14	AC155786	AC155786 Papio anu	C 450	37	80.4	227496	14	AC133763	AC133763 Rattus no
C 378	37	80.4	163388	15	AP005199	AP005199 Oryza sat	C 451	37	80.4	227720	14	AC163467	AC163467 Bos tauru
C 379	37	80.4	164328	5	AC147835	AC147835 Xenopus t	C 452	37	80.4	228365	14	AC163467	AC163467 Bos tauru
C 380	37	80.4	165350	14	AC074958	AC074958 Rattus no	C 453	37	80.4	228495	14	AC123471	AC123471 Rattus no
C 381	37	80.4	166083	8	AC007782	AC007782 Homo sapi	C 454	37	80.4	230910	1	AE017179	AE017179 Porphyrom
C 382	37	80.4	166347	8	AC098810	AC098810 Papio anu	C 455	37	80.4	231059	9	AC121497	AC121497 Mus muscu
C 383	37	80.4	166420	14	AC012112	AC012112 Homo sapi	C 456	37	80.4	231946	14	AC115544	AC115544 Rattus no
C 384	37	80.4	166642	14	AC009543	AC009543 Homo sapi	C 457	37	80.4	234549	14	AC095245	AC095245 Rattus no
C 385	37	80.4	169568	14	AC135064	AC135064 Homo sapi	C 458	37	80.4	234750	14	AC112811	AC112811 Rattus no
C 386	37	80.4	169912	9	AC121996	AC121996 Mus muscu	C 459	37	80.4	235977	14	AC132507	AC132507 Rattus no
C 387	37	80.4	170632	8	AL136321	AL136321 Human DNA	C 460	37	80.4	236254	14	AC132526	AC132526 Rattus no
C 388	37	80.4	171745	8	AC093663	AC093663 Homo sapi	C 461	37	80.4	237878	14	AC106542	AC106542 Rattus no
C 389	37	80.4	172462	14	AC162118	AC162118 Cercopith	C 462	37	80.4	238822	5	CR385066	CR385066 Zebrafish
C 390	37	80.4	172576	8	CNS01DSG	AL121775 Human chr	C 463	37	80.4	238887	14	AC109745	AC109745 Rattus no
C 391	37	80.4	175375	14	AC162073	AC162073 Bos tauru	C 464	37	80.4	239240	14	AC094155	AC094155 Rattus no
C 392	37	80.4	177862	8	AC068812	AC068812 Homo sapi	C 465	37	80.4	239312	14	AC112572	AC112572 Rattus no
C 393	37	80.4	178252	14	AL356386	AL356386 Homo sapi	C 466	37	80.4	239704	8	AC093791	AC093791 Homo sapi
C 394	37	80.4	180817	14	AL350559	AC150559 Medicago	C 467	37	80.4	240971	14	AC120956	AC120956 Rattus no
C 395	37	80.4	181061	8	AC115282	AC115282 Homo sapi	C 468	37	80.4	241139	14	AC118398	AC118398 Rattus no
C 396	37	80.4	182639	8	AC104307	AC104307 Homo sapi	C 469	37	80.4	241139	14	AC137443	AC137443 Rattus no
C 397	37	80.4	182805	5	CR392001	CR392001 Zebrafish	C 470	37	80.4	242377	14	AC128635	AC128635 Rattus no
C 398	37	80.4	183425	8	AC097064	AC097064 Homo sapi	C 471	37	80.4	242561	14	AC099408	AC099408 Pan trogl
C 399	37	80.4	184254	8	AC138391	AC138391 Homo sapi	C 472	37	80.4	242730	14	AC103302	AC103302 Rattus no
C 400	37	80.4	184511	8	AC009032	AC009032 Homo sapi	C 473	37	80.4	243043	14	AC103302	AC103302 Rattus no
C 401	37	80.4	184875	9	AC159460	AC159460 Mus muscu	C 474	37	80.4	243289	14	AC128764	AC128764 Rattus no
C 402	37	80.4	184943	9	BX294183	AC19460 Mouse DNA	C 475	37	80.4	243470	14	AC110647	AC110647 Rattus no
C 403	37	80.4	185437	15	AP005422	AP005422 Oryza sat	C 476	37	80.4	243553	14	AC096320	AC096320 Rattus no
C 404	37	80.4	185576	14	AC025691	AC025691 Homo sapi	C 477	37	80.4	243731	14	AC161660	AC161660 Bos tauru
C 405	37	80.4	188165	14	AC151074	AC151074 Bos tauru	C 478	37	80.4	244277	14	AC094248	AC094248 Rattus no
C 406	37	80.4	189080	14	BX649407	BX649407 Danio rer	C 479	37	80.4	244281	14	AC125891	AC125891 Rattus no
C 407	37	80.4	189752	8	AC161596	AC161596 Mus muscu	C 480	37	80.4	245297	14	AC126625	AC126625 Rattus no
C 408	37	80.4	190185	8	AC026464	AC026464 Homo sapi	C 481	37	80.4	245484	14	AC094714	AC094714 Rattus no
C 409	37	80.4	190439	14	AC073741	AC073741 Mus muscu	C 482	37	80.4	245552	14	AC128420	AC128420 Rattus no
C 410	37	80.4	194320	9	AC090659	AC090659 Mus muscu	C 483	37	80.4	247345	14	AC157231	AC157231 Bos tauru
C 411	37	80.4	195115	9	AC158910	AC158910 Mus muscu	C 484	37	80.4	247831	14	AC098072	AC098072 Rattus no
C 412	37	80.4	196665	9	AC141560	AC141560 Mus muscu	C 485	37	80.4	248036	14	AC155043	AC155043 Bos tauru
C 413	37	80.4	197144	8	AC073347	AC073347 Homo sapi	C 486	37	80.4	249629	9	AC113720	AC113720 Rattus no
C 414	37	80.4	197172	8	AC146315	AC146315 Otlemur	C 487	37	80.4	251417	14	AC106252	AC106252 Rattus no
C 415	37	80.4	198338	9	AC087098	AC087098 Genomic s	C 488	37	80.4	254630	14	AC108269	AC108269 Rattus no
C 416	37	80.4	198551	9	AC068605	AC068605 Mus muscu	C 489	37	80.4	256221	14	AL356371	AL356371 Homo sapi
C 417	37	80.4	198768	15	AY678298	AY678298 Lycopersi	C 490	37	80.4	256983	14	AC106285	AC106285 Rattus no
C 418	37	80.4	199337	14	CT025523	CT025523 Mus muscu	C 491	37	80.4	258653	14	AC151154	AC151154 Bos tauru
C 419	37	80.4	200624	9	AL592065	AL592065 Mouse DNA	C 492	37	80.4	259824	14	AC098374	AC098374 Rattus no
C 420	37	80.4	201133	9	AL713915	AL713915 Mouse DNA	C 493	37	80.4	260126	14	AC120589	AC120589 Rattus no
C 421	37	80.4	202902	14	AC128020	AC128020 Rattus no	C 494	37	80.4	261024	1	AE017256	AE017256 Wolbachia
C 422	37	80.4	203982	8	AC067854	AC067854 Homo sapi	C 495	37	80.4	261730	14	AC127779	AC127779 Rattus no
C 423	37	80.4	207702	9	AC060781	AC060781 Mus muscu	C 496	37	80.4	262038	5	BX511227	BX511227 Zebrafish
C 424	37	80.4	207736	14	AC155670	AC155670 Bos tauru	C 497	37	80.4	264504	1	AE017041	AE017041 Bacillus
C 425	37	80.4	208028	9	AC126260	AC126260 Mus muscu	C 498	37	80.4	265155	14	AC097116	AC097116 Rattus no
C 426	37	80.4	210228	14	AC151379	AC151379 Callithri	C 499	37	80.4	265663	14	AC106285	AC106285 Rattus no
C 427	37	80.4	210430	14	AC128771	AC128771 Rattus no	C 500	37	80.4	273602	14	AE016067	AE016067 Bos tauru
C 428	37	80.4	211088	9	AC131661	AC131661 Mus muscu	C 501	37	80.4	276034	1	AE017015	AE017015 Bacillus
C 429	37	80.4	211950	14	AC164880	AC164880 Homo sapi	C 502	37	80.4	279831	14	AC158809	AC158809 Bos tauru
C 430	37	80.4	212070	14	AC108315	AC108315 Rattus no	C 503	37	80.4	284897	14	AE016757	AE016757 Escherich
C 431	37	80.4	213072	14	AC154861	AC154861 Mus muscu	C 504	37	80.4	300413	1	AE017054	AE017054 Oryza sat
C 432	37	80.4	213217	14	AC140252	AC140252 Mus muscu	C 505	37	80.4	302085	15	AE017054	AE017054 Oryza sat
C 433	37	80.4	215044	14	AC156235	AC156235 Bos tauru	C 506	37	80.4	309457	14	AC160280	AC160280 Bos tauru
C 434	37	80.4	215581	14	AC106216	AC106216 Rattus no	C 507	37	80.4	321052	1	AE017281	AE017281 Bacillus
C 435	37	80.4	216194	9	AC096051	AC096051 Rattus no	C 508	37	80.4	340000	8	HS21C009	HS21C009 Homo sapi
C 436	37	80.4	219399	8	AC072062	AC072062 Homo sapi	C 509	37	80.4	352	10	BV332298	BV332298 S230P6456
C 437	37	80.4	220541	14	AC166827	AC166827 Mus muscu	C 510	37	80.4	359	6	AR203655	AR203655 Sequence
C 438	37	80.4	221353	14	AC097170	AC097170 Rattus no	C 511	37	80.4	359	6	AR305890	AR305890 Sequence
C 439	37	80.4	221526	9	AC117668	AC117668 Mus muscu	C 512	37	80.4	390	6	CQ752269	CQ752269 Sequence
C 440	37	80.4	221548	9	AC115689	AC115689 Mus muscu	C 513	37	80.4	450	13	VSVRNAPRS	VSVRNAPRS Vesicular s
C 441	37	80.4	221748	9	AC079216	AC079216 Mus muscu	C 514	37	80.4	517	10	G84867	G84867 S208P6389FF
							C 514	37	80.4	528	1	AY912485	AY912485 Pseudomon

C 515	36	78.3	585	2	AY500864	AY500864	Pennaeus m	588	36	78.3	9448	6	BD094913	BD094913	G protein
C 516	36	78.3	585	10	G96028	G96028	S209P601ARA	589	36	78.3	9448	8	AB042411	AB042411	Homo sapi
C 517	36	78.3	602	10	BV034351	BV034351	S212P6022	C 590	36	78.3	9581	1	PAE223604	PAE223604	Pseudomon
C 518	36	78.3	650	10	BV392987	BV392987	S243P6205	C 591	36	78.3	12339	9	AB104852	AB104852	Pseudomon
C 519	36	78.3	661	8	HSIGK7	Z00005	Human germ	592	36	78.3	14773	9	AB084238	AB084238	Mus muscu
C 520	36	78.3	661	8	HUMY201B11	DO061191	Pseudomon	593	36	78.3	15722	2	AF003132	AF003132	Caenorhab
C 521	36	78.3	664	1	AB162949	AB162949	Serratia	C 594	36	78.3	16578	1	D90721	D90721	Escherichia
C 522	36	78.3	718	1	AB162949	AB162949	Serratia	C 595	36	78.3	16885	8	AC091949	AC091949	Homo sapi
C 523	36	78.3	719	1	AB162948	AB162948	Serratia	C 596	36	78.3	18325	9	AL844211	AL844211	Mouse DNA
C 524	36	78.3	724	6	BD267205	BD267205	Compositi	C 597	36	78.3	19714	1	D90722	D90722	Escherichia
C 525	36	78.3	724	6	AR566726	AR566726	Sequence	598	36	78.3	20035	1	AE008711	AE008711	Salmonell
C 526	36	78.3	739	1	AB162947	AB162947	Serratia	599	36	78.3	24709	2	CRS26H9A	CRS26H9A	Caenorhabdi
C 527	36	78.3	739	1	AB162950	AB162950	Serratia	600	36	78.3	36143	2	U41013	U41013	Caenorhabdi
C 528	36	78.3	740	8	HSIGKLO6	X71894	H. sapiens g	601	36	78.3	36676	8	HSJ858B16	HSJ858B16	Human DNA
C 529	36	78.3	741	1	AB040994	AB040994	Serratia	C 602	36	78.3	39002	8	AF001219	AF001219	Homo sapi
C 530	36	78.3	741	1	AB195637	AB195637	Pseudomon	C 603	36	78.3	39494	8	AF001217	AF001217	Homo sapi
C 531	36	78.3	741	1	AB195638	AB195638	Achromoba	C 604	36	78.3	39946	8	AF001220	AF001220	Homo sapi
C 532	36	78.3	741	1	D78375	D78375	Pseudomonas	605	36	78.3	43241	8	HSJ9613	HSJ9613	Homo sapi
C 533	36	78.3	741	1	AY251052	AY251052	Pseudomon	606	36	78.3	4582	14	AC012831	AC012831	Drosophil
C 534	36	78.3	741	6	AX110700	AX110700	Sequence	C 607	36	78.3	48860	1	AF184956	AF184956	Bacillus
C 535	36	78.3	743	10	BV600581	BV600581	S217P6045	608	36	78.3	48249	8	AC131147	AC131147	Homo sapi
C 536	36	78.3	769	10	BV475105	BV475105	G591P6311	609	36	78.3	49105	9	AL731716	AL731716	Mouse DNA
C 537	36	78.3	783	10	BV615910	BV615910	S217P6051	610	36	78.3	53203	15	AP006683	AP006683	Lotus cor
C 538	36	78.3	880	1	AB074433	AB074433	Pseudomon	C 611	36	78.3	59476	14	AC101069	AC101069	Mus muscu
C 539	36	78.3	880	1	AB074434	AB074434	Pseudomon	C 612	36	78.3	61712	15	AB017061	AB017061	Arabidops
C 540	36	78.3	880	1	AB074435	AB074435	Achromoba	613	36	78.3	63002	14	AC129533	AC129533	Mus muscu
C 541	36	78.3	1000	10	CNS061BY	AL400100	T3 end of	614	36	78.3	68244	14	AC101375	AC101375	Mus muscu
C 542	36	78.3	1090	1	AY625689	AY625689	Pseudomon	C 615	36	78.3	68702	8	AL445193	AL445193	Human DNA
C 543	36	78.3	1119	5	AY874346	AY874346	Xenopus f	616	36	78.3	73427	8	HS32419	HS32419	Human DNA
C 544	36	78.3	1119	6	AR556739	AR556739	Sequence	C 617	36	78.3	76887	14	AC007775	AC007775	Homo sapi
C 545	36	78.3	1135	5	AY874348	AY874348	Xenopus b	618	36	78.3	77781	6	AX818190	AX818190	Sequence
C 546	36	78.3	1141	5	AY874347	AY874347	Xenopus r	619	36	78.3	77781	6	HS1409	HS1409	Human DNA
C 547	36	78.3	1141	5	AY874350	AY874350	Xenopus l	620	36	78.3	86196	8	AL592046	AL592046	Human DNA
C 548	36	78.3	1142	5	AY874344	AY874344	Xenopus w	621	36	78.3	88186	14	AC138810	AC138810	Homo sapi
C 549	36	78.3	1225	1	AY625688	AY625688	Pseudomon	622	36	78.3	92000	9	AF133300	AF133300	Mus muscu
C 550	36	78.3	1286	1	S71932	S71932	blaIMP-meta	C 623	36	78.3	93377	14	CR962138	CR962138	Medicago
C 551	36	78.3	1354	1	KPNRDK4	D29636	Klebsiella	C 624	36	78.3	95219	9	AL607073	AL607073	Mouse DNA
C 552	36	78.3	1386	6	AX496961	AX496961	Sequence	C 625	36	78.3	95908	8	AC008628	AC008628	Homo sapi
C 553	36	78.3	1386	6	AX703470	AX703470	Sequence	C 626	36	78.3	99855	15	OSJN00175	OSJN00175	Oryza sat
C 554	36	78.3	1386	6	AX720910	AX720910	Sequence	C 627	36	78.3	100000	8	AP000069	AP000069	Homo sapi
C 555	36	78.3	1386	6	AX925299	AX925299	Sequence	C 628	36	78.3	100298	14	AP007517	AP007517	Lotus cor
C 556	36	78.3	1386	6	AX925558	AX925558	Sequence	C 629	36	78.3	100364	8	AC007590	AC007590	Homo sapi
C 557	36	78.3	1609	9	BC086595	BC086595	Rattus no	C 630	36	78.3	100927	8	AL157828	AL157828	Human DNA
C 558	36	78.3	1670	9	BC018354	BC018354	Mus muscu	C 631	36	78.3	101386	8	AL359551	AL359551	Human DNA
C 559	36	78.3	1689	9	BC059831	BC059831	Mus muscu	C 632	36	78.3	101693	14	AC151349	AC151349	Xenopus t
C 560	36	78.3	1745	9	BC086601	BC086601	Rattus no	C 633	36	78.3	103419	14	AL139221	AL139221	Homo sapi
C 561	36	78.3	1884	6	AX496957	AX496957	Sequence	C 634	36	78.3	106179	1	CP000020	CP000020	28
C 562	36	78.3	1884	6	AX703466	AX703466	Sequence	C 635	36	78.3	108210	14	AL158845	AL158845	Homo sapi
C 563	36	78.3	1884	6	AX720906	AX720906	Sequence	C 636	36	78.3	110000	1	AE005174	AE005174	09
C 564	36	78.3	1884	6	AX925295	AX925295	Sequence	C 637	36	78.3	110000	1	AE005174	AE005174	10
C 565	36	78.3	1884	6	AX925554	AX925554	Sequence	C 638	36	78.3	110000	1	AE005674	AE005674	08
C 566	36	78.3	1967	8	HSIGKLO16	X71890	H. sapiens g	C 639	36	78.3	110000	1	CP000099	CP000099	13
C 567	36	78.3	2236	15	AK067882	AK067882	Oryza sat	C 640	36	78.3	110000	1	U00096	U00096	08
C 568	36	78.3	2271	15	AK070853	AK070853	Oryza sat	C 641	36	78.3	110000	1	AE001720	AE001720	04
C 569	36	78.3	2278	2	AB122063	AB122063	Crassost	C 642	36	78.3	110000	1	BA000007	BA000007	09
C 570	36	78.3	2331	15	AB121016	AB121016	Oryza sat	C 643	36	78.3	110000	1	CP000026	CP000026	24
C 571	36	78.3	2442	1	D50438	D50438	Serratia ma	C 644	36	78.3	110000	1	CP000033	CP000033	17
C 572	36	78.3	2445	15	BT000177	BT000177	Arabidops	C 645	36	78.3	110000	1	CP000058	CP000058	45
C 573	36	78.3	2612	15	AY136382	AY136382	Arabidops	C 646	36	78.3	110000	14	AC145943	AC145943	Gallus ga
C 574	36	78.3	2798	2	LEIDITRA	M85072	Leishmania	C 647	36	78.3	110000	14	TANN3	TANN3	13
C 575	36	78.3	3292	1	PACATAAC6	X98393	Pseudomonas	648	36	78.3	110000	15	AP008214	AP008214	079
C 576	36	78.3	3434	13	GPU34761	U34761	Goose parvo	C 649	36	78.3	110000	15	AP008214	AP008214	276
C 577	36	78.3	3517	6	BD094908	BD094908	G protein	C 650	36	78.3	110000	15	AP008217	AP008217	035
C 578	36	78.3	3517	6	CS104061	CS104061	Sequence	651	36	78.3	110000	15	AP008217	AP008217	035
C 579	36	78.3	3517	6	AX549355	AX549355	Sequence	652	36	78.3	110000	15	CR382122	CR382122	09
C 580	36	78.3	3517	8	AB042410	AB042410	Homo sapi	653	36	78.3	110000	15	CR382122	CR382122	10
C 581	36	78.3	3911	1	AB188812	AB188812	Pseudomon	654	36	78.3	110000	15	CR382138	CR382138	18
C 582	36	78.3	3940	1	AP416297	AP416297	Serratia	C 655	36	78.3	110000	15	AE017348	AE017348	07
C 583	36	78.3	4049	15	AK066916	AK066916	Oryza sat	C 656	36	78.3	110000	15	AP008209	AP008209	044
C 584	36	78.3	4415	1	AJ640197	AJ640197	Acinetoba	C 657	36	78.3	110000	15	AP008209	AP008209	077
C 585	36	78.3	5106	13	GPU25749	U25749	Goose parvo	C 658	36	78.3	110000	15	AP008210	AP008210	219
C 586	36	78.3	6287	1	STYRESM	M90544	Salmonella	C 659	36	78.3	110000	15	AP008210	AP008210	220
C 587	36	78.3	7534	1	AB070224	AB070224	Serratia	660	36	78.3	110000	15	AP008211	AP008211	041

C 661	36	78.3	110000	15	AP008211_269	Continuation (270	734	36	78.3	155676	14	AC044780	AC044780 Homo sapi
C 662	36	78.3	110000	15	AP008211_284	Continuation (285	735	36	78.3	155690	9	AC121885	AC121885 Mus muscu
C 663	36	78.3	110000	15	AP008213_152	Continuation (153	736	36	78.3	156342	9	AC158566	AC158566 Mus muscu
C 664	36	78.3	110627	14	AP008031	AP008031 Lotus cor	C 737	36	78.3	156745	15	AC164523	AC164523 Mus muscu
C 665	36	78.3	111107	8	AC005924	AC005924 Homo sapi	738	36	78.3	156772	15	AC087553	AC087553 Oryza sat
C 666	36	78.3	111370	8	AC067815	AC067815 Homo sapi	C 739	36	78.3	156835	14	AC068059	AC068059 Homo sapi
C 667	36	78.3	112261	8	AL139338	AL139338 Human DNA	740	36	78.3	158629	9	AC156400	AC156400 Mus muscu
C 668	36	78.3	116530	14	AC166147	AC166147 Medicago	C 741	36	78.3	159082	9	AC124408	AC124408 Mus muscu
C 669	36	78.3	116788	8	AC007247	AC007247 Homo sapi	742	36	78.3	159387	8	HS455H14	AL023280 Human DNA
C 670	36	78.3	117930	8	AC0073271	AC0073271 Homo sapi	743	36	78.3	159501	8	AC108117	AC108117 Homo sapi
C 671	36	78.3	118812	9	AL929446	AL929446 Mouse DNA	744	36	78.3	159691	8	AC025160	AC025160 Homo sapi
C 672	36	78.3	119500	15	AC098832	AC098832 Oryza sat	745	36	78.3	160487	8	AC084058	AC084058 Homo sapi
C 673	36	78.3	119909	15	AC104278	AC104278 Oryza sat	C 746	36	78.3	160666	9	AC123805	AC123805 Mus muscu
C 674	36	78.3	120557	8	AC004955	AC004955 Homo sapi	747	36	78.3	160869	14	AC147815	AC147815 Xenopus t
C 675	36	78.3	120579	5	BX323579	BX323579 Zebrafish	C 748	36	78.3	160990	8	AL157702	AL157702 Human DNA
C 676	36	78.3	120700	8	AL606845	AL606845 Human DNA	749	36	78.3	161081	14	AC022997	AC022997 Homo sapi
C 677	36	78.3	121130	15	OSJN00292	BX548156 Oryza sat	C 750	36	78.3	161990	14	AC024723	AC024723 Homo sapi
C 678	36	78.3	122538	8	AC011942	AC011942 Homo sapi	C 751	36	78.3	162098	14	CR628321	CR628321 Danio rer
C 679	36	78.3	122715	9	AL845335	AL845335 Mouse DNA	752	36	78.3	162467	9	AC114622	AC114622 Mus muscu
C 680	36	78.3	123849	15	AP004044	AP004044 Oryza sat	753	36	78.3	163120	8	AC011454	AC011454 Homo sapi
C 681	36	78.3	124231	14	AC016466	AC016466 Homo sapi	C 754	36	78.3	163317	14	AC142170	AC142170 Rattus no
C 682	36	78.3	124652	14	AC157374	AC157374 Medicago	755	36	78.3	163650	8	AL445932	AL445932 Human DNA
C 683	36	78.3	125243	15	AC128643	AC128643 Oryza sat	C 756	36	78.3	163868	5	CR792429	CR792429 Zebrafish
C 684	36	78.3	126181	8	AC074131	AC074131 Homo sapi	C 757	36	78.3	164125	8	AC018693	AC018693 Homo sapi
C 685	36	78.3	128398	8	AC004817	AC004817 Homo sapi	C 758	36	78.3	164736	14	AL390855	AL390855 Homo sapi
C 686	36	78.3	129548	14	AC141022	AC141022 Rattus no	C 759	36	78.3	165326	8	AL159974	AL159974 Human DNA
C 687	36	78.3	129593	13	AF512031	AF512031 Chorlaton	C 760	36	78.3	167092	5	CR376773	CR376773 Zebrafish
C 688	36	78.3	129992	8	AL354897	AL354897 Human DNA	761	36	78.3	167108	8	CNS01RHA	AL161666 Human chr
C 689	36	78.3	130146	15	AP005251	AP005251 Oryza sat	C 762	36	78.3	167320	5	BX004996	BX004996 Zebrafish
C 690	36	78.3	131694	8	AC012321	AC012321 Homo sapi	763	36	78.3	167799	8	AC0100798	AC0100798 Homo sapi
C 691	36	78.3	132372	15	AC146631	AC146631 Medicago	764	36	78.3	167891	8	AC013439	AC013439 Homo sapi
C 692	36	78.3	132894	15	AC141323	AC141323 Medicago	C 765	36	78.3	168693	14	AC151198	AC151198 Bos tauru
C 693	36	78.3	133797	14	AC105740	AC105740 Sus scrof	766	36	78.3	169781	9	AF336378	AF336378 Mus muscu
C 694	36	78.3	134362	14	AC084853	AC084853 Homo sapi	C 767	36	78.3	169842	8	AF301237	AF301237 Homo sapi
C 695	36	78.3	135365	14	AC157960	AC157960 Strongylo	C 768	36	78.3	169913	14	AC031980	AC031980 Homo sapi
C 696	36	78.3	136384	15	AP004559	AP004559 Oryza sat	C 769	36	78.3	170088	8	AC026123	AC026123 Homo sapi
C 697	36	78.3	136713	15	AC084817	AC084817 Oryza sat	770	36	78.3	170312	14	AC113225	AC113225 Rattus no
C 698	36	78.3	136932	8	AC068035	AC068035 Homo sapi	C 771	36	78.3	171226	8	AC099670	AC099670 Homo sapi
C 699	36	78.3	136943	8	AL359842	AL359842 Human DNA	C 772	36	78.3	171442	15	AC125471	AC125471 Oryza sat
C 700	36	78.3	137391	8	AL161732	AL161732 Human DNA	C 773	36	78.3	172029	2	AC092494	AC092494 Drosophil
C 701	36	78.3	139022	14	AC136171	AC136171 Rattus no	774	36	78.3	172334	8	AC010650	AC010650 Homo sapi
C 702	36	78.3	139961	15	AP004668	AP004668 Oryza sat	C 775	36	78.3	172367	8	AC016822	AC016822 Homo sapi
C 703	36	78.3	140594	5	BX510911	BX510911 Zebrafish	C 776	36	78.3	172465	8	AC110080	AC110080 Homo sapi
C 704	36	78.3	141160	9	AL645970	AL645970 Mouse DNA	777	36	78.3	172625	14	AC024596	AC024596 Homo sapi
C 705	36	78.3	141754	5	BX890619	BX890619 Zebrafish	C 778	36	78.3	172893	14	AC024596	AC024596 Homo sapi
C 706	36	78.3	142746	9	AC155927	AC155927 Mus muscu	C 779	36	78.3	173290	14	AC151123	AC151123 Bos tauru
C 707	36	78.3	142755	14	AC155419	AC155419 Zee maye	C 780	36	78.3	173589	9	AC109276	AC109276 Mus muscu
C 708	36	78.3	143028	14	CR749173	CR749173 Danio rer	781	36	78.3	173726	9	AC116685	AC116685 Mus muscu
C 709	36	78.3	143028	14	CR749173	CR749173 Danio rer	782	36	78.3	173988	9	AC138112	AC138112 Mus muscu
C 710	36	78.3	143257	9	AC115848	AC115848 Mus muscu	783	36	78.3	174210	15	AC123897	AC123897 Oryza sat
C 711	36	78.3	143506	8	AC092833	AC092833 Homo sapi	784	36	78.3	174480	5	AL928975	AL928975 Zebrafish
C 712	36	78.3	143769	8	AC004914	AC004914 Homo sapi	785	36	78.3	174564	14	AC121704	AC121704 Rattus no
C 713	36	78.3	143790	8	AC055811	AC055811 Homo sapi	C 786	36	78.3	175092	9	AC122491	AC122491 Mus muscu
C 714	36	78.3	144108	9	BX936355	BX936355 Mouse DNA	787	36	78.3	175326	8	CNS01DTQ	AL122838 Human chr
C 715	36	78.3	144446	8	AC146219	AC146219 Pan trogl	C 788	36	78.3	175928	9	AC115871	AC115871 Mus muscu
C 716	36	78.3	145253	8	AC006329	AC006329 Homo sapi	C 789	36	78.3	176188	9	AC124574	AC124574 Mus muscu
C 717	36	78.3	146100	14	AC159565	AC159565 Otolenur	C 790	36	78.3	176409	9	AC102130	AC102130 Mus muscu
C 718	36	78.3	146930	15	AC118672	AC118672 Genomic s	C 791	36	78.3	176613	14	CR847801	CR847801 Danio rer
C 719	36	78.3	146939	8	AC092945	AC092945 Homo sapi	C 792	36	78.3	176814	14	AC121405	AC121405 Rattus no
C 720	36	78.3	147890	15	AP004395	AP004395 Oryza sat	C 793	36	78.3	177109	14	AC118161	AC118161 Rattus no
C 721	36	78.3	148902	14	AC105741	AC105741 Sus scrof	794	36	78.3	177314	9	AC022209	AC022209 Homo sapi
C 722	36	78.3	149254	9	AC132223	AC132223 Mus muscu	795	36	78.3	177600	9	AC124452	AC124452 Mus muscu
C 723	36	78.3	149543	14	CT009489	CT009489 Mus muscu	C 796	36	78.3	177832	14	AC151489	AC151489 Callichri
C 724	36	78.3	151076	5	BX005194	BX005194 Zebrafish	C 797	36	78.3	178058	8	AC009139	AC009139 Homo sapi
C 725	36	78.3	151405	14	AC011172	AC011172 Homo sapi	798	36	78.3	178074	5	BX539324	BX539324 Zebrafish
C 726	36	78.3	151643	9	AC127227	AC127227 Mus muscu	C 799	36	78.3	178348	14	AC132931	AC132931 Mus muscu
C 727	36	78.3	151691	5	AC145785	AC145785 Xenopus t	C 800	36	78.3	178779	9	AC124409	AC124409 Mus muscu
C 728	36	78.3	153248	14	AC112500	AC112500 Homo sapi	801	36	78.3	179081	9	AC126274	AC126274 Mus muscu
C 729	36	78.3	153624	8	HS1018K9	AL031726 Human DNA	802	36	78.3	179593	14	AC155946	AC155946 Mus muscu
C 730	36	78.3	154859	8	AC012596	AC012596 Homo sapi	803	36	78.3	179115	5	BX323449	BX323449 Zebrafish
C 731	36	78.3	154893	8	AC068711	AC068711 Homo sapi	C 804	36	78.3	179427	14	AC147584	AC147584 Otolenur
C 732	36	78.3	155580	8	AC024597	AC024597 Homo sapi	805	36	78.3	179557	8	AC151479	AC151479 Pan trogl
C 733	36	78.3	155580	8	AC024946	AC024946 Homo sapi	806	36	78.3	179652	14	AC166177	AC166177 Mus muscu

807	36	78.3	180363	9	AC154755	AC154755 Mus muscu	c 880	36	78.3	208660	14	AC020577	AC020577 Homo sapi
c 808	36	78.3	180637	9	AL844491	AL844491 Mouse DNA	881	36	78.3	208801	9	AC150560	AC150560 Mus muscu
809	36	78.3	180835	14	AC128108	AC128108 Rattus no	882	36	78.3	208862	14	AC161269	AC161269 Mus muscu
810	36	78.3	181519	9	AC112968	AC112968 Mus muscu	883	36	78.3	209670	9	AC125409	AC125409 Mus muscu
811	36	78.3	181873	14	AC150834	AC150834 Callithri	c 884	36	78.3	209957	9	AC158236	AC158236 Mus muscu
c 812	36	78.3	182105	5	AL844179	AL844179 Zebrafish	c 885	36	78.3	210005	14	AC128869	AC128869 Mus muscu
c 813	36	78.3	182324	14	AC102276	AC102276 Mus muscu	c 886	36	78.3	210123	14	CR388419	CR388419 Danio rer
c 814	36	78.3	182507	8	BS000596	BS000596 Pan trogl	887	36	78.3	210163	9	AL645524	AL645524 Mouse DNA
c 815	36	78.3	182859	14	AL845341	AL845341 Mus muscu	c 888	36	78.3	210202	14	AC124427	AC124427 Mus muscu
c 816	36	78.3	183192	9	AY016021	AY016021 Mus muscu	889	36	78.3	211201	8	AC009958	AC009958 Homo sapi
c 817	36	78.3	183349	9	AC107794	AC107794 Mus muscu	890	36	78.3	211750	8	AP006306	AP006306 Homo sapi
c 818	36	78.3	183679	8	AC017095	AC017095 Homo sapi	891	36	78.3	212599	5	AL929222	AL929222 Zebrafish
819	36	78.3	183689	14	AC146129	AC146129 Pan trogl	892	36	78.3	215147	5	AC145944	AC145944 Gallus ga
820	36	78.3	183773	9	AC154213	AC154213 Mus muscu	893	36	78.3	215156	14	AC095447	AC095447 Rattus no
c 821	36	78.3	184872	9	AL672243	AL672243 Mouse DNA	894	36	78.3	215283	14	AC154905	AC154905 Mus muscu
822	36	78.3	184906	14	BX914208	BX914208 Danio rer	895	36	78.3	215819	14	AC099288	AC099288 Rattus no
823	36	78.3	185754	14	AC160972	AC160972 Mus muscu	c 896	36	78.3	215993	14	AC153043	AC153043 Bos tauru
824	36	78.3	185879	14	AC109019	AC109019 Rattus no	c 897	36	78.3	216102	14	AC112314	AC112314 Rattus no
c 825	36	78.3	186675	5	AL845306	AL845306 Zebrafish	c 898	36	78.3	216255	14	AC118841	AC118841 Rattus no
826	36	78.3	186734	9	EX005053	EX005053 Mouse DNA	899	36	78.3	216638	14	AC149735	AC149735 Bos tauru
c 827	36	78.3	187074	14	CR848675	CR848675 Danio rer	900	36	78.3	217376	14	AL603630	AL603630 Mus muscu
828	36	78.3	187532	14	AC006333	AC006333 Homo sapi	c 901	36	78.3	217376	14	AL603630	AL603630 Mus muscu
829	36	78.3	187712	14	AC120815	AC120815 Rattus no	902	36	78.3	218675	14	AC015847	AC015847 Homo sapi
830	36	78.3	188416	14	AC140097	AC140097 Sus scrof	903	36	78.3	219356	14	AC157592	AC157592 Mus muscu
c 831	36	78.3	188701	14	AC102515	AC102515 Mus muscu	c 904	36	78.3	219491	14	AC092256	AC092256 Mus muscu
832	36	78.3	188862	14	AC159740	AC159740 Mus muscu	905	36	78.3	219502	14	AC164312	AC164312 Mus muscu
833	36	78.3	189019	5	BX640507	BX640507 Zebrafish	c 906	36	78.3	220260	9	AC164436	AC164436 Mus muscu
c 834	36	78.3	189427	14	AC134945	AC134945 Rattus no	907	36	78.3	221335	14	CR936838	CR936838 Bos tauru
c 835	36	78.3	190264	14	AC161624	AC161624 Pan trogl	908	36	78.3	221952	14	AC150582	AC150582 Bos tauru
836	36	78.3	190375	14	AC109933	AC109933 Rattus no	c 909	36	78.3	222059	9	AC127336	AC127336 Mus muscu
c 837	36	78.3	190375	14	AC109933	AC109933 Rattus no	c 910	36	78.3	222059	5	BX465868	BX465868 Zebrafish
c 838	36	78.3	190960	14	AC153882	AC153882 Mus muscu	911	36	78.3	222343	14	AC114531	AC114531 Rattus no
839	36	78.3	191671	5	BX511000	BX511000 Zebrafish	912	36	78.3	223037	9	AC161606	AC161606 Mus muscu
c 840	36	78.3	192010	9	AL732320	AL732320 Mouse DNA	913	36	78.3	223171	14	AC162608	AC162608 Bos tauru
841	36	78.3	192387	8	CNS01DVO	AL133444 Human chr	c 914	36	78.3	224843	14	AC095744	AC095744 Rattus no
842	36	78.3	193005	9	AC110917	AC110917 Mus muscu	915	36	78.3	225185	14	AC107007	AC107007 Rattus no
843	36	78.3	193917	9	AL844153	AL844153 Mouse DNA	c 916	36	78.3	226001	15	AP331808	AP331808 Zea mays
844	36	78.3	194032	14	AC019150	AC019150 Homo sapi	c 917	36	78.3	226843	14	AC157187	AC157187 Bos tauru
845	36	78.3	194060	9	AC125349	AC125349 Mus muscu	c 918	36	78.3	227010	14	AC157187	AC157187 Gallus ga
846	36	78.3	194881	9	AC087890	AC087890 Mus muscu	c 919	36	78.3	227044	9	AC107774	AC107774 Mus muscu
c 847	36	78.3	195113	14	AC161510	AC161510 Mus muscu	920	36	78.3	227469	14	AC097678	AC097678 Rattus no
c 848	36	78.3	195144	8	AC087292	AC087292 Homo sapi	921	36	78.3	228481	9	AC153579	AC153579 Mus muscu
849	36	78.3	195311	9	AC121585	AC121585 Mus muscu	c 922	36	78.3	228845	14	AC115377	AC115377 Rattus no
850	36	78.3	195396	5	AL935148	AL935148 Zebrafish	c 923	36	78.3	228924	14	AC126701	AC126701 Rattus no
c 851	36	78.3	195505	5	BX537287	BX537287 Zebrafish	924	36	78.3	229372	14	AC134003	AC134003 Rattus no
852	36	78.3	195556	14	AC018382	AC018382 Homo sapi	925	36	78.3	230011	14	AC126712	AC126712 Rattus no
c 853	36	78.3	197271	14	AC112105	AC112105 Rattus no	c 926	36	78.3	230114	14	AC153036	AC153036 Bos tauru
c 854	36	78.3	197872	14	BX897666	BX897666 Danio rer	c 927	36	78.3	230988	9	AC115749	AC115749 Mus muscu
c 855	36	78.3	197891	14	AC109837	AC109837 Rattus no	928	36	78.3	231926	9	AL662779	AL662779 Mouse DNA
c 856	36	78.3	199500	9	AL596103	AL596103 Mouse DNA	929	36	78.3	232058	14	AC112024	AC112024 Rattus no
857	36	78.3	199980	9	AC127310	AC127310 Mus muscu	930	36	78.3	232977	9	AC153877	AC153877 Mus muscu
c 858	36	78.3	200409	5	BX469922	BX469922 Zebrafish	c 931	36	78.3	233448	14	AC073692	AC073692 Mus muscu
859	36	78.3	200815	14	AC137450	AC137450 Rattus no	932	36	78.3	233783	9	AC118698	AC118698 Mus muscu
c 860	36	78.3	201354	9	AC122819	AC122819 Mus muscu	933	36	78.3	233916	14	AC136533	AC136533 Rattus no
861	36	78.3	201747	14	AC163876	AC163876 Bos tauru	934	36	78.3	234084	9	AC095825	AC095825 Rattus no
c 862	36	78.3	201977	9	AC147564	AC147564 Mus muscu	935	36	78.3	234133	14	AC095750	AC095750 Rattus no
c 863	36	78.3	202010	14	CR388072	CR388072 Danio rer	c 936	36	78.3	234563	14	AC123093	AC123093 Rattus no
c 864	36	78.3	202012	9	AC130530	AC130530 Mus muscu	937	36	78.3	234847	14	AC111336	AC111336 Rattus no
865	36	78.3	202540	14	AC022971	AC022971 Homo sapi	938	36	78.3	235044	14	AC163749	AC163749 Mus muscu
c 866	36	78.3	203897	9	AC121596	AC121596 Mus muscu	c 939	36	78.3	235453	14	AC131525	AC131525 Rattus no
c 867	36	78.3	204131	14	AC128898	AC128898 Rattus no	c 940	36	78.3	235516	14	AC163072	AC163072 Bos tauru
c 868	36	78.3	204420	14	AC162624	AC162624 Bos tauru	c 941	36	78.3	235929	14	AC106086	AC106086 Rattus no
869	36	78.3	204453	9	AC125539	AC125539 Mus muscu	c 942	36	78.3	236060	14	AC130954	AC130954 Rattus no
c 870	36	78.3	204534	14	AC084400	AC084400 Mus muscu	c 943	36	78.3	237143	14	AC099253	AC099253 Rattus no
871	36	78.3	204730	5	BX890575	BX890575 Zebrafish	c 944	36	78.3	237317	14	AC097420	AC097420 Rattus no
c 872	36	78.3	204954	14	AC128028	AC128028 Rattus no	c 945	36	78.3	237339	14	AC113666	AC113666 Rattus no
873	36	78.3	205208	8	AC147213	AC147213 Pan trogl	c 946	36	78.3	237542	14	AC134123	AC134123 Rattus no
874	36	78.3	205318	9	AC122894	AC122894 Mus muscu	947	36	78.3	238887	14	AC097850	AC097850 Rattus no
875	36	78.3	205831	14	AC074192	AC074192 Homo sapi	c 948	36	78.3	238937	14	AC106652	AC106652 Rattus no
c 876	36	78.3	206466	8	AC010548	AC010548 Homo sapi	c 949	36	78.3	239385	14	AC109686	AC109686 Rattus no
c 877	36	78.3	206852	8	AC091182	AC091182 Homo sapi	c 950	36	78.3	239746	14	AC115387	AC115387 Rattus no
c 878	36	78.3	207287	14	BX546497	BX546497 Mus muscu	c 951	36	78.3	239746	14	AC125557	AC125557 Rattus no
879	36	78.3	208005	14	AC146293	AC146293 Saimiri b	952	36	78.3	240094	14	AC163357	AC163357 Mus muscu

```

953 36 78.3 240507 14 AC103202
954 36 78.3 242056 14 AC096280
955 36 78.3 242077 14 AC095805
956 36 78.3 242697 14 AC094159
957 36 78.3 243362 14 AC096234
958 36 78.3 244571 14 AC128600
959 36 78.3 245365 14 AC107404
960 36 78.3 246896 14 AC105498
961 36 78.3 247265 14 AC120685
962 36 78.3 249287 14 AC142252
963 36 78.3 250275 14 AC095695
964 36 78.3 250404 14 AC128509
965 36 78.3 250404 14 AC162352
966 36 78.3 250600 14 AC136646
967 36 78.3 250743 14 AC097042
968 36 78.3 251086 14 AC099142
969 36 78.3 251141 14 AC116235
970 36 78.3 251322 14 AC157234
971 36 78.3 251976 14 AC111885
972 36 78.3 252108 9 AC138642
973 36 78.3 252787 14 AC154971
974 36 78.3 253097 14 AC096445
975 36 78.3 253237 14 AC106474
976 36 78.3 253738 9 AC091478
977 36 78.3 254436 14 AC103569
978 36 78.3 254831 14 AC157027
979 36 78.3 256760 14 AC094580
980 36 78.3 257047 14 AC090879
981 36 78.3 257608 14 AC153408
982 36 78.3 258938 14 AC099416
983 36 78.3 259415 14 AC156178
984 36 78.3 259450 14 AC094112
985 36 78.3 262108 14 AC105555
986 36 78.3 262453 14 AC164535
987 36 78.3 262794 14 AC119702
988 36 78.3 264121 14 AC109097
989 36 78.3 265913 14 AC133620
990 36 78.3 267267 14 AC131350
991 36 78.3 268050 14 AC627266
992 36 78.3 268828 14 AC091513
993 36 78.3 271861 14 AC108541
994 36 78.3 273541 14 AC096596
995 36 78.3 274480 14 AC115513
996 36 78.3 276343 14 AC125570
997 36 78.3 280005 14 AC119021
998 36 78.3 280337 14 AC122969
999 36 78.3 283336 14 AC097749
1000 36 78.3 286332 14 AC022105

```

ALIGNMENTS

```

RESULT 1
AX829164 927 bp DNA linear PAT 12-DEC-2003
LOCUS
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE
1 Mack,D.H., Gish,K.C. and Afar,D.
METHODS Methods of diagnosis of breast cancer, compositions and methods of
screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)
FEATURES
source
1. .927
/organism="Homo sapiens"

```

```

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.22 Length: 927
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-12 (1-9) x AX829164 (1-927)
Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 283 AACCTCTTCTTACCGGCAACAGCTG 309
RESULT 2
AX467373 1260 bp DNA linear PAT 16-JUL-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 Myers,K., Drury,N. and Carroll,M.
AUTHORS
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. .1260
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"
ORIGIN
Alignment Scores:
Pred. No.: 3.02 Length: 1260
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-12 (1-9) x AX467373 (1-1260)
Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 280 AACCTCTTCTTACCGGCAATCAGCTG 306
RESULT 3
AX821533 1260 bp DNA linear PAT 10-DEC-2003
LOCUS
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM Felis catus (cat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 Carroll,M.M., Kingman,S.M. and Redchenko,I.M.
AUTHORS MHC class I peptide epitopes from the human St4 tumor-associated
antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;

```

```

FEATURES             Oxford Biomedica (UK) Limited (GB)
source              Location/Qualifiers
                    1..1260
                    /organism="Felis catus"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9685"

ORIGIN
Alignment Scores:      Length: 1260
Pred. No.:            Matches: 9
Score:                46.00
Percent Similarity:    100.0%
Best Local Similarity: 100.0%
Query Match:          100.0%
Indels:               0
Gaps:                 0
DB:                   0

US-10-774-176-12 (1-9) x AX821533 (1-1260)

QY      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      280 AACCTCTTCCTCACCGGCAATCAGCTG 306

RESULT 4
LOCUS   AX821548                1260 bp      DNA      linear      PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION   AX821548.1 GI:39724930
KEYWORDS .
SOURCE   Felis catus (cat)
ORGANISM Felis catus
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Rutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
          Felinae; Felis.
REFERENCE 1
AUTHORS  Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
TITLE    MHC class II peptide epitope of 5t4 antigen
JOURNAL  Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)

FEATURES             Location/Qualifiers
source              1..1260
                    /organism="Felis catus"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9685"

ORIGIN
Alignment Scores:      Length: 1260
Pred. No.:            Matches: 9
Score:                46.00
Percent Similarity:    100.0%
Best Local Similarity: 100.0%
Query Match:          100.0%
Indels:               0
Gaps:                 0
DB:                   0

US-10-774-176-12 (1-9) x AX821548 (1-1260)

QY      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      280 AACCTCTTCCTCACCGGCAATCAGCTG 306

RESULT 5
BD249731
LOCUS   BD249731                1263 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION   BD249731.1 GI:33059501
KEYWORDS  JP 2002530060-A/1.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
          Mammalia; Rutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.
REFERENCE 1 (bases 1 to 1263)

```

```

AUTHORS             Carroll,M.W. and Myers,K.A.
TITLE               Polypeptide
JOURNAL             Patent: JP 2002530060-A 1 17-SEP-2002;
                    OXFORD BIOMEDICA LTD
COMMENT             OS Homo sapiens (human)
                    PN JP 2002530060-A/1
                    PD 17-SEP-2002
                    PF 18-NOV-1999 JP 2000582415
                    PR 18-NOV-1998 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
                    30-JUL-1999 GB 9917995.4
                    PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
                    PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
                    C07K19/00,
                    PC C12N15/00,
                    CC Polypeptide
                    FT Key
                    FT source 1..1263
                    Location/Qualifiers
                    /organism="Homo sapiens (human)"
                    /db_xref="taxon:9606"

FEATURES             Location/Qualifiers
source              1..1263
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      Length: 1263
Pred. No.:            Matches: 9
Score:                46.00
Percent Similarity:    100.0%
Best Local Similarity: 100.0%
Query Match:          100.0%
Indels:               0
Gaps:                 0
DB:                   0

US-10-774-176-12 (1-9) x BD249731 (1-1263)

QY      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      283 AACCTCTTCCTTACCGGCAACAGCTG 309

RESULT 6
LOCUS   AX025011                1263 bp      DNA      linear      PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO029428.
ACCESSION AX025011
VERSION   AX025011.1 GI:10184932
KEYWORDS .
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Rutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.
REFERENCE 1
AUTHORS  Carroll,M.W. and Myers,K.A.
TITLE    Polypeptide
JOURNAL  Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)

FEATURES             Location/Qualifiers
source              1..1263
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      Length: 1263
Pred. No.:            Matches: 9
Score:                46.00
Percent Similarity:    100.0%
Best Local Similarity: 100.0%
Query Match:          100.0%
Indels:               0
Gaps:                 0
DB:                   0

US-10-774-176-12 (1-9) x AX025011 (1-1263)

```

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 Db 283 AACCTCTTCTTACCGGCAACACGCTG 309

RESULT 7
 LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
 DEFINITION Sequence 14 from Patent WO0136486.
 ACCESSION AX149553
 VERSION AX149553.1 GI:14347991
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.

REFERENCE 1
 AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
 Ellard,F.M. and Myers,K.A.
 TITLE Antibodies
 JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1. .1263
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="514"

ORIGIN
 Alignment Scores:
 Pred. No.: 3.02 Length: 1263
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x AX149553 (1-1263)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 Db 283 AACCTCTTCTTACCGGCAACACGCTG 309

RESULT 8
 LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
 DEFINITION Sequence 1 from Patent EP1160323.
 ACCESSION AX316086
 VERSION AX316086.1 GI:17899278
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE 1
 AUTHORS Carroll,M.W. and Myers,K.A.
 TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
 JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1. .1263
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 3.02 Length: 1263
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0

DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x AX316086 (1-1263)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 Db 283 AACCTCTTCTTACCGGCAACACGCTG 309

RESULT 9
 LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
 DEFINITION Sequence 1 from Patent WO0238612.
 ACCESSION AX467371
 VERSION AX467371.1 GI:21900602
 KEYWORDS
 SOURCE Canis sp.
 ORGANISM Canis sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.

REFERENCE 1
 AUTHORS Myers,K., Drury,N. and Carroll,M.
 TITLE Polypeptide
 JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1. .1263
 /organism="Canis sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:9616"

ORIGIN
 Alignment Scores:
 Pred. No.: 3.02 Length: 1263
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x AX467371 (1-1263)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 |||||
 Db 283 AACCTCTTCTTACCGGCAACACGCTG 309

RESULT 10
 LOCUS CQ731678 2053 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 17612 from Patent WO02068579.
 ACCESSION CQ731678
 VERSION CQ731678.1 GI:42308932
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE 1
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 17612 06-SEP-2002;
 PE Corporation (NY) (US)

FEATURES
 source
 1. .2053
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:

```
Pred. No.: 4.9 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x CQ731678 (1-2053)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 367 AACCTCTTCCTTACCGCAACGCTG 393

RESULT 11
H5T40A H5T40A 2053 bp RNA linear PRI 18-APR-2005
LOCUS Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.
DEFINITION VERSION 229083
ACCESSION 229083.1 GI:435654
KEYWORDS 5T4 gene; 5T4 oncofoetal antigen.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2053)
AUTHORS Myers,K.A., Rahi-Saund,V., Davison,M.D., Young,J.A., Cheater,A.J.
and Stern,P.L.
TITLE Isolation of a cDNA encoding 5T4 oncofoetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
J. Biol. Chem. 269 (12), 9319-9324 (1994)
PUBMED 8132670
REFERENCE 2 (bases 1 to 2053)
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
FEATURES
source
1. .2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62. .372
/product="LRR N-terminal flank"
/label=N-Flank
85. .1347
/codon_start=1
/evidence=experimental
/product="5T4 oncofoetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translat="MPGCGSRGPAAGDGLRLARLALVLLGWSSSPTSSASSFSSS
APFLASVAPQPLPDQPCALCESEARTVKVNRNLTEVPTDLPAYVRNLFLTGQ
LAVLPACAFARRPEPLAELANLNSGSLRDEVRAGAFELPSLROLDI.SHPPLADLSP
AFGNSASVAPSPPLVRLILNHIIVPPDERQNRSEFGVVAALLAGRALQRLRLA
SNFLYLPRLVLAQLSLRLDLSNLSLVTVVSPRNLTHLSLHLDNALKLHNG
TLAELQGLPHIRVFLDNPMVCDKMDVMTLKBETVWGCKDLTCAYPEKMNRYL
LELNSADLDCDPLIPPSLQTSYVFLGIVLALIGALFLLVLYLNRKGIKKWMHNRDCA
RDMGEGHYRYEINADPRNLNLSNSDV"
130. .171
/sig_peptide
373. .966
/misc_RNA
/product="Leucine rich repeat region"
/label=LRRs

misc_RNA 966. .1119
/product="LRR C-terminal flank"
/label=C-flank
1153. .1215
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

ORIGIN
Alignment Scores:
Pred. No.: 4.9 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x H55T40A (1-2053)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 367 AACCTCTTCCTTACCGCAACGCTG 393

RESULT 12
BD127282
LOCUS BD127282 2359 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127282
VERSION BD127282.1 GI:23222227
KEYWORDS JP 2002017375-A/2713.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Negai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2713 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2713
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10, C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS
Location/Qualifiers
(424) ..(1572).
source
1. .2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5.63 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
```


US-10-774-176-12 (1-9) x BD127282 (1-2359)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||

Db 706 AACCTCTCTTACCGGACACGCTG 732

RESULT 13

CQ782724 LOCUS 2359 bp DNA linear PAT 17-MAR-2004

DEFINITION Sequence 2864 from Patent EP1396543.

ACCESSION CQ782724

VERSION CQ782724.1 GI:45502667

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1 Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.

TITLE Primers for synthesizing full length cDNA clones and their use

JOURNAL Patent: EP 1396543-A 2864 10-MAR-2004;
Research Association for Biotechnology (JP)

FEATURES Location/Qualifiers

1..2359
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"

CDS

APFLAGVAVSQPQLDQPCALCESEARTVKVNRNLTEVPTDLPAYVRNLFTGNG
LAVLPAGAPAPPLPGLAALNLGSRGLDEVRGAPPHLPRLDLSHNPDLADLSPP
AFSGNSVASRPPLELLIINHVPPDERQNSFEGWVAALLAGRALGLRLLEIA
SNHFLYPRDVLQAPLRLHLSNLSISITVYSFRLTHLSLHLEHNAKLVLENG
TLAELQGLPHIRVFLDNPWCDCMDMTWLTKEVTVQGGKRLTCAYPEKMRNRL
LELSADLDCDPLPPSLQTSYVPLGIVLALIGALFLLVLYLRKGIKK

ORIGIN

Alignment Scores:
Pred. No.: 5.63 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x CQ782724 (1-2359)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||

Db 706 AACCTCTCTTACCGGACACGCTG 732

RESULT 14

AK074786 LOCUS 2359 bp mRNA linear PRI 03-SEP-2002

DEFINITION Homo sapiens cDNA FLJ30305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.

ACCESSION AK074786

VERSION AK074786.1 GI:22760460

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1

AUTHORS Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,

TITLE

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2359)

AUTHORS Isogai, T. and Otsuki, T.

TITLE Direct Submission

JOURNAL

Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

Location/Qualifiers

1..2359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 5.63 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x AK074786 (1-2359)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||

Db 706 AACCTCTCTTACCGGACACGCTG 732

RESULT 15

BD127283 LOCUS 2361 bp DNA linear PAT 18-SEP-2002

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD127283

VERSION BD127283.1 GI:23222228

KEYWORDS JP 2002017375-A/2714.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1 (bases 1 to 2361)

AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002017375-A 2714 22-JAN-2002;

COMMENT HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002017375-A/2714

PD 22-JAN-2002

PF 07-JUL-2000 JP 2000253172

PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO

PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
 SHINICHI KOJIMA,
 PI TETSUJI OTSUKI, HISASHI KOGA
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
 10, C12P21/02, C12Q1/68//C12P21/08, G06P17/30, C12N15/00, C12N5/00 CC
 PC C12P21/02, C12Q1/68//C12P21/08, G06P17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers
 (426)..(1695).

FEATURES

source
 1..2361
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 5.63 Length: 2361
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x BD127283 (1-2361)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

|||||
 Db 708 AACCTCTTCTTACCGGCAACGAGCTG 734

RESULT 16

CQ782726
 LOCUS 2361 bp DNA linear PAT 17-MAR-2004
 DEFINITION Sequence 2866 from Patent EP1396543.
 CQ782726
 ACCESSION
 VERSION CQ782726.1 GI:45502669

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE

AUTHORS Oka, T., Nishikawa, T., Isogai, T., Hayashi, K., Iehii, S., Kawai, Y.,
 Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
 Koga, H.
 TITLE Primers for synthesizing full length cDNA clones and their use
 JOURNAL Patent: EP 1396543-A 2866 10-MAR-2004;
 Research Association for Biotechnology (JP)

FEATURES

source
 1..2361
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 426..1688
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAF85961.1"
 /db_xref="GI:45502670"

CDS

|||||
 Db 708 AACCTCTTCTTACCGGCAACGAGCTG 734

ORIGIN

Alignment Scores:
 Pred. No.: 5.63 Length: 2361
 Score: 46.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x CQ782726 (1-2361)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

|||||
 Db 708 AACCTCTTCTTACCGGCAACGAGCTG 734

RESULT 17

AX961916
 LOCUS 2361 bp DNA linear PAT 14-JAN-2004
 DEFINITION Sequence 127 from Patent WO03104277.
 AX961916
 ACCESSION
 VERSION AX961916.1 GI:40881326

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE

AUTHORS Sugahara, T., Mateuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.

JOURNAL Stat6 activation gene

Patent: WO 03104277-A 127 18-DEC-2003;

Asahi Kasei Kabushiki Kaisha (JP)

FEATURES

source
 1..2361
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 426..1688
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAF06467.1"
 /db_xref="GI:40881327"

CDS

|||||
 Db 708 AACCTCTTCTTACCGGCAACGAGCTG 734

ORIGIN

Alignment Scores:
 Pred. No.: 5.63 Length: 2361
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x AX961916 (1-2361)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

|||||
 Db 708 AACCTCTTCTTACCGGCAACGAGCTG 734

RESULT 18

AK074790
 LOCUS 2361 bp mRNA linear PRI 09-JUL-2005
 DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
 to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.

ACCESSION

AK074790

VERSION

AK074790.1 GI:22760466

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 2361)
REFERENCE
AUTHORS
Otsuki, T., Ota, T., Nishikawa, T., Hayashi, K., Suzuki, Y., Yamamoto, J., Makamatsu, A., Kimura, K., Sakamoto, K., Hatano, N., Kawai, Y., Ishii, S., Saito, K., Kojima, S., Sugiyama, T., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Nagai, K., Sugano, S. and Isogai, T.
TITLE
Signal Sequence and Keyword Trap in silico for Selection of Full-Length Human cDNAs Encoding Secretion or Membrane Proteins from Oligo-Capped cDNA Libraries
JOURNAL
DNA Res. 12, 117-126 (2005)
2
REFERENCE
AUTHORS
Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hio, Y., Saito, K., Yamamoto, J., Makamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuh, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Ninomiya, K.
TITLE
NEDO human cDNA sequencing project
JOURNAL
Unpublished
3 (bases 1 to 2361)
REFERENCE
AUTHORS
Isogai, T. and Otsuki, T.
TITLE
Direct Submission
JOURNAL
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).
LOCATION/Qualifiers
1. .2361
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/notes="cloning vector: pME18SPL3
mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

ORIGIN
Alignment Scores:
Pred. No.: 5.63 Length: 2361
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x AK074790 (1-2361)
QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||
DB 708 AACCTCTTCTTACCGGACACAGCTG 734
RESULT 19
BC037161
LOCUS
DEFINITION
Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317 IMAGE:4138906), complete cds.
ACCESSION
BC037161 GI:33872201
VERSION
BC037161
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 2379)
REFERENCE
AUTHORS
Krausner, R.D., Collins, P.S., Wagner, L.H., Derge, J.G., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J.J., Haieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Shapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2379)
REFERENCE
AUTHORS
Strausberg, R.
TITLE
Direct Submission
JOURNAL
Submitted (03-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK
COMMENT
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:22713382.
Contact: MGC help desk
Email: cgapbe-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghig, P., Hansen, N., Ho, S.-L., Karlins, B., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 26 Row: m Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5729717.
LOCATION/Qualifiers
1. .2379
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:15317 IMAGE:4138906"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/note="vector: pOTB7"
1. .2379
/gene="TPBG"
/note="synonyms: M6P1, 5T4-AG, 5T4"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"

```

CDS
427..1689
/genes="TPBG"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="AAH37161.1"
/db_xref="GI:22713383"
/db_xref="GeneID:7162"
/db_xref="MIM:130920"
/translation="MPGCGSRGPAAGDGLRLARLARLALVLGWSSSPTSSASPSSS
APFLASAVSQPPLDQPCALCESEARTVKCVNRNLTEVTDLPAYVRLFLTGNO
LAVPLGAFARPPPLAEALNLSGRLDEVRAGAFELPSLRQLDLSHNPADLSPP
AFSGSNASGAPSLVELILNHIIVPPDEQRNRSFEGMVAAALAGRALQGLRLLELA
SNHFLYPLRDVLAQLPSRLHLDLSNLSVLTYSFRLNTHLESLEDNALKVLHNG
TLAELOGLPHIRVFLDNNPWCDCMADVTMLKETVVGOKPLTCTAYPEKWRNVL
LELNSADLDCDPLPSPLOTSYVFLGIVLALIGALFLVLVLYLRKGIKKWMMHNRDAC
RDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN
Alignment Scores:
Pred. No.: 5.68 Length: 2379
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x BC037161 (1-2379)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||
DB 709 AACCTCTCTCTTACCGGCACACGCTG 735

RESULT 20
AB168308 2714 bp mRNA linear PRI 18-JUN-2005
LOCUS Macaca fascicularis testis cDNA clone: QtsA-11109, similar to human
DEFINITION trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.
ACCESSION AB168308
VERSION AB168308.1 GI:67967899
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE
1 International consortium for macaque cDNA sequencing and analysis.
  DNA sequences of macaque genes expressed in brain or testis and its
  evolutionary implications
  Unpublished
JOURNAL
AUTHORS Osada, N., Hirata, M., Tanuma, R., Kusuda, J., Hida, M., Suzuki, Y.,
  Sugano, S., Gojobori, T., Shen, J.C.-K., Wu, C.I. and Hashimoto, K.
TITLE Substitution rate and structural divergence of 5'UTR evolution:
  Comparative analysis between human and cynomolgus monkey cDNAs
  Unpublished
JOURNAL
REFERENCE 3 (bases 1 to 2714)
AUTHORS Hashimoto, K., Kusuda, J. and Sugano, S.
TITLE Direct Submission
JOURNAL
SUBMITTED (18-MAR-2004) Katsuyuki Hashimoto, National Institute of
  Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
  1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan
  (E-mail: khashim@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
  Tel: 81-3-5285-1111 (ex.2120), Fax: 81-3-5285-1181)
  The International consortium for macaque cDNA sequencing and
  analysis consists of: Department of Virology and Human Genome
  Center, Institute of Medical Science, The University of Tokyo,
  Tokyo, Japan; Division of Genetic Resources, National Institute of
  Infectious Diseases of Japan, Tokyo, Japan; National Health
  Research Institute, Taipei, Taiwan; Institute of Molecular Biology,
  Academia Sinica, Taipei, Taiwan; Department of Ecology & Evolution,
  University of Chicago, Chicago, IL, USA; Center for Information
  Biology, National Institute of Genetics of Japan, Mishima, Japan.

```

```

Clone distribution: clone distribution information can be found at:
http://www.nih.go.jp/yoken/genebank/
Lab host: TOP10
Vector: pME188-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACTGTGTG)
R. Site2: DraIII (CACCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGTCGCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME188-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by oligo-capping method. Libraries were made from:
QCC6: cerebellum cortex
QMPA: parietal lobe
QTPA: temporal lobe right
QFLA: frontal lobe left
QMOA: medulla oblongata
QBSA: brain stem
QORA: occipital lobe right
QTSa: testis
Custom primers were used for 5' and 3'-end sequencing. The
full-insert sequencing was done by primer-walking method using ABI
DNA sequencer.
FEATURES
Location/Qualifiers
1..2714
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="QtsA-11109"
/sex="male"
/clone_lib="macaque cDNA library QtsA"
/dev_stage="adult"
764..2026
/notes="unnamed protein product; Homo sapiens trophoblast
glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3"
/protein_id="BAB00432.1"
/db_xref="GI:67967900"
/translation="MPGCGSRGPAAGDGLRLARLARLALVLGWSSSPTSSASSSSS
APFLASASAPPLDQPCALCESEARTVKCVNRNLTEVTDLPAYVRLFLTGNO
LAVPLGAFARPPPLAEALNLSGRLDEVRAGAFELPSLRQLDLSHNPADLSPP
AFSGSNASGAPSLVELILNHIIVPPDDQRNRSFEGMVAAALVAGRALQGLHLELA
SNHFLYPLRDVLAQLPSRLHLDLSNLSVLTYSFRLNTHLESLEDNALKVLHNG
TLAELOGLPHIRVFLDNNPWCDCMADVTMLKETVVGOKPLTCTAYPEKWRNVL
LELNSADLDCDPLPSPLOTSYVFLGIVLALIGALFLVLVLYLRKGIKKWMMHNRDAC
RDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN
Alignment Scores:
Pred. No.: 6.47 Length: 2714
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x AB168308 (1-2714)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||
DB 1046 AACCTCTCTCTTACCGGCACACGCTG 1072

RESULT 21
HSA012159
LOCUS Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
DEFINITION Homo sapiens 5T4 oncofetal trophoblast glycoprotein.
ACCESSION AJ012159
VERSION AJ012159.1 GI:3805946
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1. King, K.W., Sheppard, P.C., Westwater, C., Stern, P.L. and Myers, K.A. Organisation of the mouse and human 5T4 oncofoetal leucine-rich glycoprotein genes and expression in foetal and adult murine tissues

Biochim. Biophys. Acta 1445 (3), 257-270 (1999)

10366710

2 (bases 1 to 5551)

Myers, K.A.

Direct Submission

Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson Institute for Cancer Research, Christie Hospital, Wilmslow Road, Manchester, M20 9BX, UK

Location/Qualifiers

1. .5551

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

2698..2703

/bound_moiety="Sp1"

2704..2709

/bound_moiety="Sp1"

2716..5400

/gene="5T4"

2716..2800

/gene="5T4"

/evidence="experimental"

2801..3092

/gene="5T4"

/evidence="experimental"

3093..5400

/gene="5T4"

/evidence="experimental"

3431..4693

/gene="5T4"

/codon_start=1

/product="5T4 oncofetal trophoblast glycoprotein"

/protein_id="CAA09930.1"

/db_xref="GI:3805947"

/db_xref="GOA:Q13641"

/db_xref="InterPro:IPR000372"

/db_xref="InterPro:IPR000483"

/db_xref="InterPro:IPR001611"

/db_xref="InterPro:IPR003591"

/db_xref="UniProt/TREMBL:Q13641"

/translation="MPGCGSRGPAAGRLRLARLALVLIGWVSSSPSTSSASSFSSS
APLASAVSAQPPPLPQPCALCSCSAARTVKCNRLTEVPTDLPAYVRNLPLTGNQ
LAVLPAGAFARRPPLAEALNLSGRLDEVRAGAEHLPSLQDLSHNPLADLSP
AFSGSNASVAPSLVELLIHNVPPEDERQNRSGFQGVVAALLAGALQGLRLLELA
SNHPLPLDRVLQPLSLHLDLSNNSLSVTSFPRNTHLESLEHEDNALVHLNG
TLAELQGLPHIRVFLDNNPWCDCNMADMTLKEVTEVQSGDRLTCAYPEKRNRLV
LELNSADLDCDPTLPSPSLQTSYVFLGIVLIGALIFLLVLYLNKRGKIKKMHNRDAD
RDHMEGVHYEINADPRLTNLSNSDV"

3431..3516

/gene="5T4"

3517..4690

/gene="5T4"

/product="5T4 oncofetal trophoblast glycoprotein"

5331..5336

/gene="5T4"

5380..5385

/gene="5T4"

sig_peptide

mat_peptide

polyA_signal

polyA_signal

ORIGIN

Alignment Scores:

Pred. No.: 13.2

Score: 46.00

Percent Similarity: 100.0%

Best Local Similarity: 100.0%

Query Match: 100.0%

Length: 5551

Matches: 9

Conservative: 0

Mismatches: 0

Indels: 0

DB:	8	Gaps:	0
US-10-774-176-12 (1-9) x HSA012159 (1-5551)			
Qy	1 AsnLeuPheLeuThrGlyAsnGlnLeu 9		
Db	3713 AACCTCTTCTTACCGGCAACACGCTG 3739		
RESULT 22			
AP008208_073			
WPCOMMENT			
Sequence split into 360 fragments LOCUS AP008208 Accession AP008208			
Fragment Name	Begin	End	
AP008208_000	1	110000	
AP008208_001	100001	210000	
AP008208_002	200001	310000	
AP008208_003	300001	410000	
AP008208_004	400001	510000	
AP008208_005	500001	610000	
AP008208_006	600001	710000	
AP008208_007	700001	810000	
AP008208_008	800001	910000	
AP008208_009	900001	1010000	
AP008208_010	1000001	1110000	
AP008208_011	1100001	1210000	
AP008208_012	1200001	1310000	
AP008208_013	1300001	1410000	
AP008208_014	1400001	1510000	
AP008208_015	1500001	1610000	
AP008208_016	1600001	1710000	
AP008208_017	1700001	1810000	
AP008208_018	1800001	1910000	
AP008208_019	1900001	2010000	
AP008208_020	2000001	2110000	
AP008208_021	2100001	2210000	
AP008208_022	2200001	2310000	
AP008208_023	2300001	2410000	
AP008208_024	2400001	2510000	
AP008208_025	2500001	2610000	
AP008208_026	2600001	2710000	
AP008208_027	2700001	2810000	
AP008208_028	2800001	2910000	
AP008208_029	2900001	3010000	
AP008208_030	3000001	3110000	
AP008208_031	3100001	3210000	
AP008208_032	3200001	3310000	
AP008208_033	3300001	3410000	
AP008208_034	3400001	3510000	
AP008208_035	3500001	3610000	
AP008208_036	3600001	3710000	
AP008208_037	3700001	3810000	
AP008208_038	3800001	3910000	
AP008208_039	3900001	4010000	
AP008208_040	4000001	4110000	
AP008208_041	4100001	4210000	
AP008208_042	4200001	4310000	
AP008208_043	4300001	4410000	
AP008208_044	4400001	4510000	
AP008208_045	4500001	4610000	
AP008208_046	4600001	4710000	
AP008208_047	4700001	4810000	
AP008208_048	4800001	4910000	
AP008208_049	4900001	5010000	
AP008208_050	5000001	5110000	
AP008208_051	5100001	5210000	
AP008208_052	5200001	5310000	
AP008208_053	5300001	5410000	
AP008208_054	5400001	5510000	
AP008208_055	5500001	5610000	
AP008208_056	5600001	5710000	
AP008208_057	5700001	5810000	
AP008208_058	5800001	5910000	
AP008208_059	5900001	6010000	

AP008208_060	6000001	6110000	AP008208_133	13300001	13410000
AP008208_061	6100001	6210000	AP008208_134	13400001	13510000
AP008208_062	6200001	6310000	AP008208_135	13500001	13610000
AP008208_063	6300001	6410000	AP008208_136	13600001	13710000
AP008208_064	6400001	6510000	AP008208_137	13700001	13810000
AP008208_065	6500001	6610000	AP008208_138	13800001	13910000
AP008208_066	6600001	6710000	AP008208_139	13900001	14010000
AP008208_067	6700001	6810000	AP008208_140	14000001	14110000
AP008208_068	6800001	6910000	AP008208_141	14100001	14210000
AP008208_069	6900001	7010000	AP008208_142	14200001	14310000
AP008208_070	7000001	7110000	AP008208_143	14300001	14410000
AP008208_071	7100001	7210000	AP008208_144	14400001	14510000
AP008208_072	7200001	7310000	AP008208_145	14500001	14610000
AP008208_073	7300001	7410000	AP008208_146	14600001	14710000
AP008208_074	7400001	7510000	AP008208_147	14700001	14810000
AP008208_075	7500001	7610000	AP008208_148	14800001	14910000
AP008208_076	7600001	7710000	AP008208_149	14900001	15010000
AP008208_077	7700001	7810000	AP008208_150	15000001	15110000
AP008208_078	7800001	7910000	AP008208_151	15100001	15210000
AP008208_079	7900001	8010000	AP008208_152	15200001	15310000
AP008208_080	8000001	8110000	AP008208_153	15300001	15410000
AP008208_081	8100001	8210000	AP008208_154	15400001	15510000
AP008208_082	8200001	8310000	AP008208_155	15500001	15610000
AP008208_083	8300001	8410000	AP008208_156	15600001	15710000
AP008208_084	8400001	8510000	AP008208_157	15700001	15810000
AP008208_085	8500001	8610000	AP008208_158	15800001	15910000
AP008208_086	8600001	8710000	AP008208_159	15900001	16010000
AP008208_087	8700001	8810000	AP008208_160	16000001	16110000
AP008208_088	8800001	8910000	AP008208_161	16100001	16210000
AP008208_089	8900001	9010000	AP008208_162	16200001	16310000
AP008208_090	9000001	9110000	AP008208_163	16300001	16410000
AP008208_091	9100001	9210000	AP008208_164	16400001	16510000
AP008208_092	9200001	9310000	AP008208_165	16500001	16610000
AP008208_093	9300001	9410000	AP008208_166	16600001	16710000
AP008208_094	9400001	9510000	AP008208_167	16700001	16810000
AP008208_095	9500001	9610000	AP008208_168	16800001	16910000
AP008208_096	9600001	9710000	AP008208_169	16900001	17010000
AP008208_097	9700001	9810000	AP008208_170	17000001	17110000
AP008208_098	9800001	9910000	AP008208_171	17100001	17210000
AP008208_099	9900001	10010000	AP008208_172	17200001	17310000
AP008208_100	10000001	10110000	AP008208_173	17300001	17410000
AP008208_101	10100001	10210000	AP008208_174	17400001	17510000
AP008208_102	10200001	10310000	AP008208_175	17500001	17610000
AP008208_103	10300001	10410000	AP008208_176	17600001	17710000
AP008208_104	10400001	10510000	AP008208_177	17700001	17810000
AP008208_105	10500001	10610000	AP008208_178	17800001	17910000
AP008208_106	10600001	10710000	AP008208_179	17900001	18010000
AP008208_107	10700001	10810000	AP008208_180	18000001	18110000
AP008208_108	10800001	10910000	AP008208_181	18100001	18210000
AP008208_109	10900001	11010000	AP008208_182	18200001	18310000
AP008208_110	11000001	11110000	AP008208_183	18300001	18410000
AP008208_111	11100001	11210000	AP008208_184	18400001	18510000
AP008208_112	11200001	11310000	AP008208_185	18500001	18610000
AP008208_113	11300001	11410000	AP008208_186	18600001	18710000
AP008208_114	11400001	11510000	AP008208_187	18700001	18810000
AP008208_115	11500001	11610000	AP008208_188	18800001	18910000
AP008208_116	11600001	11710000	AP008208_189	18900001	19010000
AP008208_117	11700001	11810000	AP008208_190	19000001	19110000
AP008208_118	11800001	11910000	AP008208_191	19100001	19210000
AP008208_119	11900001	12010000	AP008208_192	19200001	19310000
AP008208_120	12000001	12110000	AP008208_193	19300001	19410000
AP008208_121	12100001	12210000	AP008208_194	19400001	19510000
AP008208_122	12200001	12310000	AP008208_195	19500001	19610000
AP008208_123	12300001	12410000	AP008208_196	19600001	19710000
AP008208_124	12400001	12510000			
AP008208_125	12500001	12610000			
AP008208_126	12600001	12710000			
AP008208_127	12700001	12810000			
AP008208_128	12800001	12910000			
AP008208_129	12900001	13010000			
AP008208_130	13000001	13110000			
AP008208_131	13100001	13210000			
AP008208_132	13200001	13310000			

Alignment Scores:
Pred. No.: 257
Score: 46.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 15
DB:
Length: 110000
Matches: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-774-176-12 (1-9) x AP008208_073 (1-110000)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db 37749 AATTTGTTTAACTGGTAACAGCTA 37775

RESULT 23

HSJ492P14

LOCUS

DEFINITION

Human DNA sequence from clone RP3-492P14 on chromosome 6q13-15
Contains a single stranded DNA binding protein pseudogene, the TPBG
gene for trophoblast glycoprotein (574-AG) and a CpG island,
complete sequence.

ACCESSION

AL121977

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Dec 15, 2000 this sequence version replaced gi:11558491.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/chr6
RP3-492P14 is from the library RPI-3 constructed by the group of
Pietar de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES

source

1. 121909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="q13-15"
/clone="RP3-492P14"
/clone_lib="RPI-3"
100

misc_feature

/note="Clone right end: RPI-93K22"

gene

complement(1004..10982)

/locus_tag="RP3-492P14.2-001"

CDS

complement(1004..10982)

/locus_tag="RP3-492P14.2-001"

/note="match: proteins: P81877 Q99LX9 Q9BWM6 Q9CYZ8 Q9D6L4
Q9F038 Q9Y4T7"

/pseudo

misc_feature

86539
/note="Clone left end: RPI-90G1"

gene

109639..116836

mRNA

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

join(109639..109916,110631..116836)

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

/product="trophoblast glycoprotein"

/note="match: ESTs: AJ149121 AA152323 AA565852 AA643734

AL544610 AW471072 AW662538 BE260089 BF306457 BF306926

BF314984 BI196133 B1562387 BM069633 BM670613

match: cDNAs: AJ420536.1 Z29083.1"

110970..112232

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

/standard_name="OTTHUMP00000016786"

/note="match: proteins: Q13641 Q9QYD9 Q9Z0L0"

/codon_start=1

/product="trophoblast glycoprotein"

/protein_id="CAI21546.1"

/db_xref="GI:56203539"

/db_xref="Gene:12004"

/db_xref="GOA:Q13641"

/db_xref="InterPro:IPR000372"

/db_xref="InterPro:IPR000483"

/db_xref="InterPro:IPR001611"

/db_xref="InterPro:IPR003591"

/db_xref="UniProt/TREMBL:Q13641"

/translation="MPGCSRGPAAGDGLRLRLARLALVLLGWSSSPTSASFSSS

APFLAGAVSAQPLPDQCPCSAARTVKVNRNLTEVPTDLPAYVRNLPLTGNQ

LAVLPAGAPARRPLAEALNLGSRDLDEVRAGAFELPSRLDLSHNPDLADLSP

AFSGSNASVAPSPPLVELILNHLVPPEDERQNSFEGMVAAALLAGRALQGLRELA

SNHFLYPRDVLAAQLSLRHLDSNNLSVITVYFNLTHLSLHLEDNALKVLHNG

TLAEQLGLPHIRVFLDNNPWVCDHMDMTWLKETEVQVKDLITCAYPEKMRNL

LENSADLDCDPIPLPSLQTSYVPLGIVLALIGAIFLLVLYLNRRKGTKKMMHNRD

RDHMGVHYRYEINADPRLTNLSNSDV"

116817..116822

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

116836

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

121909

/note="Clone right end: RP3-492P14"

ORIGIN

Alignment Scores:

Pred. No.:	284	Length:	121909
Score:	46.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-774-176-12 (1-9) x HSJ492P14 (1-121909)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

|||||

Db 111252 AACCTCTTCTTACCGGCAACAGCTG 111278

RESULT 24

AP005756

LOCUS

DEFINITION

AP005756
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone: OSJNB0035N08.

ACCESSION

AP005756

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;

REFERENCE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.	mRNA	/gene="OSJNBb0035N08.3" complement(join(<11805. .12148,12855. .>13042))
AUTHORS	Sasaki,T., Matsumoto,T. and Katayose,Y.	misc_feature	/gene="OSJNBb0035N08.3" /note="supported by full-length cDNA(s): AK111070"
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC clone:OSJNBb0035N08		/complement(11805. .13042) /gene="OSJNBb0035N08.3" /note="contains full-length cDNA(s): AK111070"
JOURNAL	Published Only in Database (2002)	gene	non-coding transcript probably inactive due to no initiation codon in CDS"
REFERENCE	2 (bases 1 to 136267)	misc_feature	complement(13420. .13668) /gene="OSJNBb0035N08.4" complement(13420. .13668) /gene="OSJNBb0035N08.4" /note="hypothetical ORF predicted by GlimmerM
AUTHORS	Sasaki,T., Matsumoto,T. and Katayose,Y.	gene	this category is not included in IRGSP standard"
TITLE	Direct Submission	misc_feature	complement(14428. .14868) /gene="OSJNBb0035N08.5" complement(<14428. .>14868) /gene="OSJNBb0035N08.5" /note="start and end point are not identified"
JOURNAL	Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Teukuba, Ibaraki 305-8602, Japan (E-mail:tsasakienias@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)	gene	complement(14428. .14868) /gene="OSJNBb0035N08.5" /note="predicted by GeneMark.hmm etc."
COMMENT	On Jun 28, 2004 this sequence version replaced gi:42627749. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), PCRNESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/cdb/glimmer/gimr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from M13rev to -21M13 of the BAC clone. This sequence of OSJNBb0035N08 clone has an overlap with P0620H05 (DDBJ: AP005394) clone at 5' end and with OJ1711.D06 (DDBJ: AP004857) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html. Location/Qualifiers 1. .136267 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="2" /clone="OSJNBb0035N08" complement(join(157. .256,518. .608,645. .687,1665. .1820)) /gene="OSJNBb0035N08.1" complement(join(157. .256,518. .608,645. .687,1665. .1820)) /gene="OSJNBb0035N08.1" /note="hypothetical ORF predicted by GENSCAN this category is not included in IRGSP standard" complement(6895. .7272) /gene="OSJNBb0035N08.2" complement(6895. .7272) /gene="OSJNBb0035N08.2" /note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" complement(11805. .13042)	misc_feature	/gene="OSJNBb0035N08.6" 15873. .16028 /gene="OSJNBb0035N08.6" /note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" complement(19965. .23497) /gene="OSJNBb0035N08.7" complement(join(19965. .20252,20345. .20405,20630. .20777,21029. .21282,22175. .22283,22950. .23497)) /gene="OSJNBb0035N08.7" /note="supported by full-length cDNA(s): AK121870" complement(join(20211. .20252,20345. .20405,20630. .20777,21029. .21282,22175. .22283,22950. .23238)) /gene="OSJNBb0035N08.7" /note="contains EST(s): AU100805(C51513),C27285(C51513) contains full-length cDNA(s): AK121870" /codon_start=1 /product="putative aux/IAA protein" /protein_id="BAD26156.1" /db_xref="GI:49388936" /translation="MGESSEMKKISGRLLGSGWGRPSDHHRHGDQEEBEKTLLELS LGLPGGWRACRDGTTTKHSTAAANAADDDGDKSMLSLGISTLVSHSQKANKN KSGPEEEHPPTATNGNALASNNNGCFQTRSPVVPVWPFRNLATSKASL ELQNGKAKAEIKRAPIKINNDGVPICRKLIDNAPDSYEKLSLAVDKLFRGLLAA QRDLPTAGAKCQOEDVAISGLIDGTGYTLVYDEGDKVLGVDPWGMFVSSVKRL RVLKTSLSLITSGRKRTAEC" complement(24133. .24630) /gene="OSJNBb0035N08.8" complement(24133. .24630) /gene="OSJNBb0035N08.8" /note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard" complement(27886. .28131) /gene="OSJNBb0035N08.9" complement(27886. .28131) /gene="OSJNBb0035N08.9" /note="hypothetical ORF predicted by GlimmerM this category is not included in IRGSP standard"


```

gene      31945..32890
          /gene="OSUNB0035N08.10"
mRNA      join(<31945..31975,32081..32221,32569..32627,
32831..>32890)
          /gene="OSUNB0035N08.10"
          /note="start and end point are not identified"
CDS       join(31945..31975,32081..32221,32569..32627,32831..32890)
          /gene="OSUNB0035N08.10"
          /note="predicted by GeneMark.hmm etc."
          /codon start=1
          /product="hypothetical protein"
          /protein_id="BAD26157.1"
          /db_xref="GI:49388937"
          /translation="MACGKNVSGDLRPLCGSTAVDQRLCAEDRPPALSLAPASP
IASILTRKYGKGRDLRSYAVVPPKLPWAVSITMEERQAGDLVSL"
          complement(33602..35288)
          /gene="OSUNB0035N08.11"
          complement(join(33602..33842,33945..34031,34799..34969,
35069..35288))
          /gene="OSUNB0035N08.11"
          /note="supported by full-length cDNA(s): AK058248"
          complement(join(33810..33842,33945..34031,34799..34969,
35069..35194))
          /gene="OSUNB0035N08.11"
          /note="contains EST(s): C99556(E20550)
contains full-length cDNA(s): AK058248"
          /codon start=1
          /product="putative 40S ribosomal protein S24"
          /protein_id="BAD26158.1"
          /db_xref="GI:49388938"
          /translation="MADTKTAPATVLTTRKFTNRLLSRKQFVLEVLHGRANVSKAD
LKEKLAKLYEKDSNCFVFKPTFHGGGKSTGFLTYDNLDAKKYEPKYLIRNGL
ATKVKSRKQMKERKRAKIRGVKTKAGDAGKKK"
          35886..36086
          /gene="OSUNB0035N08.12"
          <35886..>36086
          /gene="OSUNB0035N08.12"
          /note="start and end point are not identified"
          35886..36086
          /gene="OSUNB0035N08.12"
          /note="predicted by FGENSEH etc."
          /codon start=1
          /product="hypothetical protein"
          /protein_id="BAD26159.1"
          /db_xref="GI:49388939"

Alignment Scores:
Pred. No.:      318      Length:      136267
Score:          46.00    Matches:      9
Percent Similarity: 100.0%  Conservative: 0
Best Local Similarity: 100.0%  Mismatches: 0
Query Match:      100.0%    Indels:      0
DB:               15       Gaps:        0

US-10-774-176-12 (1-9) x AP005756 (1-136267)

Qy      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      113557 AATTGTTTTTAACGTGTAACACGCTA 113583

RESULT 25
BD249732 1281 bp      DNA      linear      PAT 17-JUL-2003
LOCUS     Polypeptide.
DEFINITION BD249732
ACCESSION BD249732.1 GI:33059502
VERSION    JP 2002530060-A/2.
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 1281)

```

```

AUTHORS    Carroll,M.W. and Myers,K.A.
TITLE      Polypeptide
PATENT:    JP 2002530060-A 2 17-SEP-2002;
           OXFORD BIOMEDICA LTD
COMMENT     OS Mus musculus (mouse)
           PN JP 2002530060-A/2
           PD 17-SEP-2002 JP 2000582415
           PR 18-NOV-1999 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
           30-JUL-1999 GB 9917995.4
           PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
           PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
           PC C07K19/00,
           CC C12N15/00
           CC Polypeptide
           PH Key
           FT source
           FT 1..1281
           Location/Qualifiers
           source
             1..1281
             /organism="Mus musculus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      8.61      Length:      1281
Score:          44.00    Matches:      8
Percent Similarity: 100.0%  Conservative: 1
Best Local Similarity: 88.9%  Mismatches: 0
Query Match:      95.7%    Indels:      0
DB:               6       Gaps:        0

US-10-774-176-12 (1-9) x BD249732 (1-1281)

Qy      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      283 AACCTTTTCTTACCGCAACGATG 309

RESULT 26
AX025012 1281 bp      DNA      linear      PAT 15-SEP-2000
LOCUS     Sequence 2 from Patent WO0029428.
DEFINITION AX025012
ACCESSION  AX025012
VERSION     AX025012.1 GI:10184933
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE  1
AUTHORS    Carroll,M.W. and Myers,K.A.
TITLE      Polypeptide
PATENT:    WO 0029428-A 2 25-MAY-2000;
           CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
           BIOMEDICA LTD (GB)
FEATURES   Location/Qualifiers
           source
             1..1281
             /organism="Mus musculus"
             /mol_type="unassigned DNA"
             /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      8.61      Length:      1281
Score:          44.00    Matches:      8
Percent Similarity: 100.0%  Conservative: 1
Best Local Similarity: 88.9%  Mismatches: 0
Query Match:      95.7%    Indels:      0
DB:               6       Gaps:        0

US-10-774-176-12 (1-9) x AX025012 (1-1281)

```

```

QY      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      283 AACCTTTCTTACCGGCACACGATG 309

RESULT 27
LOCUS   AX316087               1281 bp      DNA      linear      PAT 14-DEC-2001
DEFINITION   Sequence 2 from Patent EP1160323.
ACCESSION   AX316087
VERSION     AX316087.1   GI:17899279
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Carroll,M.W. and Myers,K.A.
TITLE       5T4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL     Patent: EP 1160323-A 2 05-DEC-2001;
            Oxford Biomedica (UK) Limited (GB)
FEATURES
            Location/Qualifiers
            1..1281
            /organism="Mus musculus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      8.61      Length:      1281
Score:          44.00     Matches:      8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match:    95.7% Indels:      0
DB:             6 Gaps:      0

US-10-774-176-12 (1-9) x AX316087 (1-1281)

QY      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      283 AACCTTTCTTACCGGCACACGATG 309

RESULT 28
LOCUS   AF063939               2333 bp      mRNA      linear      ROD 01-JAN-2000
DEFINITION   Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA,
            complete cds.
ACCESSION   AF063939
VERSION     AF063939.1   GI:6650211
KEYWORDS
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 2333)
AUTHORS     Ninkina,N.N. and Buchman,V.L.
TITLE       Structure and expression of the rat 5T4 gene
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 2333)
AUTHORS     Buchman,V.L.
TITLE       Direct Submision
JOURNAL     Submitted (06-MAY-1998) School of Biomedical Sciences, University
            of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
            UK
FEATURES
            Location/Qualifiers
            1..2333
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /tissue_type="cerebellum"
            /dev_stage="newborn"

Gene
1..2333
/gene="5T4"
5'UTR
1..363
/gene="5T4"
364..1644
/gene="5T4"
CDS
/codon_start=1
/product="5T4 oncofetal antigen homolog"
/protein_id="AAF21770.1"
/db_xref="GI:6650212"
/translation="MPGAGSRGSPSAGDGRLLRLARLALVLLGWVSASAPSSSLPSSSTS
PAFLASGSAQPPPAERCPAECSEARTKVCNRLNLEVPADLPVYRNILFLTGNO
MTVLPAQAPARQPLADLAVNLGNHLKVEGAGAFHPLGLRLRLDLSNPLTSLSAF
TFAGSNVSTPSPLEILLNHIIVPPQORONGSPGVMVAFEGMVAALRSLALRGL
HHLASNHFYLYLPRLDLLDQPLSKHLDRNNSLVSYAFRNITHTLESLENDAL
KVLHNSTLAEWQGLHVRVFLDNNPVCDFGLMADVMVSKETFEVVPDKARLTCAPPEK
MNRNGLDLTSSDLDCDQATLPSQTSYVFLGIVLALIGALFLVLLYLNKRGIKKMMH
NIRDACRDHMEGYHYRVEINADPSLTNLSNSDV"
1645..2333
/gene="5T4"
2315..2320
/gene="5T4"

3'UTR
polyA_signal

ORIGIN
Alignment Scores:
Pred. No.:      15.6      Length:      2333
Score:          44.00     Matches:      8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match:    95.7% Indels:      0
DB:             9 Gaps:      0

US-10-774-176-12 (1-9) x AF063939 (1-2333)

QY      1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
Db      646 AACCTTTCTTCACTGGCACACGATG 672

RESULT 29
LOCUS   BC087011               2361 bp      mRNA      linear      ROD 13-DEC-2004
DEFINITION   Rattus norvegicus trophoblast glycoprotein, mRNA (cdna clone
            MGC:93332 IMAGE:7193411), complete cds.
ACCESSION   BC087011
VERSION     BC087011.1   GI:56268819
KEYWORDS
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1 (bases 1 to 2361)
AUTHORS     Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,I., Wang,J., Hsieh,P.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Heltón,B., Kettman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shrivchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smaluk,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

```

12477932
 2 (bases 1 to 2361)
 DIRECTOR MGC Project.
 DIRECT SUBMISSION
 Submitted (02-DEC-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Howard Jacobs

cDNA Library Preparation: Express Genomics

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mdpaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 186 Row: 0 Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

source

1. .2361

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="MGC:93332 IMAGE:7193411"

/tissue_type="Heart, rat (Brown Norway)"

/clone_lib="NIH_MGC_234"

/lab_host="DH10B"

/notes="Vector: pExpress1"

1. .2361

/gene="Tpbp"

/note="synonym: 5T4"

/db_xref="GeneID:83684"

/db_xref="RGD:621453"

364. .1644

/gene="Tpbp"

/codon_start=1

/product="tpbg protein"

/protein_id="AAH87011.1"

/db_xref="GI:56268820"

/db_xref="GeneID:83684"

/db_xref="RGD:621453"

/translation="MPGAGSGPSAGDGLRLARLALVLGWSASAPSSSLPSSSTS

PAALASGSAPPPARCPAACSCSEARVTKCVNRNLLVLPADLPYVRNLFUTGNQ

MTVLPAGAPQPLADLAVNLISGNHKEVGAGAFHLPGLRRDLDSHPLNLISAF

TFPAGNSVSTPSPLELLIHLHTVPEDQRQNGSFEQVAFEGKVAALRSGLALRGL

HLLELASNLYPLRDLIDQLPSLKHLDLRNLSLVTSYASFRNLTHLSLHLEDNAL

KVLNNTLAEGQLAHVRLFDNNPNVCDYMDVMVSLKETEVPVDPKARLTCAFPKK

MNRGLDLTSLDLCDAITPQSLQTSYVFLGIVLALIGALFLLVLYNRKGIKKWHI

NTRDARDHMEGYHYRYEINAPRLTNLSNSDV"

ORIGIN

Alignment Scores:

Pred. No.: 15.8 Length: 2361

Score: 44.00 Matches: 8

Percent Similarity: 100.0% Conservative: 1

Best Local Similarity: 88.9% Mismatches: 0

Query Match: 95.7% Indels: 0

DB: 9 Gaps: 0

US-10-774-176-12 (1-9) x BC087011 (1-2361)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

|||||

Db 646 AACCTTTTCTCACTGGCACACAGATG 672

RESULT 30

BC058198

LOCUS

DEFINITION

BC058198

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 2423)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, P.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2423)

Strausberg, R.

Direct Submission

Submitted (15-SEP-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Contact: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nigr.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Detrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, R., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,

Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 123 Row: p Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

source

1. .2423

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

```
/db_xref="taxon:10090"
/clone="MGC:68145 IMAGE:5353871"
/tissue_type="Mammary tumor; C3 (1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP Mam6"
/lab_host="DH10B"
/note="Vector; pCMV-SPORT6"
1. .2423
/genes="tpbg"
/note="synonym: 574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
402. .1682
/genes="tpbg"
/codon_start=1
/product="tpbg protein"
/protein_id="AAH58198.1"
/db_xref="GI:34849574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
/translation="MPGAGSRGSPAGDGRRLRLRLALVLLGWVSASAPSSVPSSSTS
PAFLASGSAQPPPAERCPAAECSEAAATKVCNRRNLLVFPADLPVYVNRNLTGNQ
MTVLPAGAFARQPPPLADLEALNLSGNHLKEVCAGAFELPGLRLDLSHNPLTNSAF
VFAGSNASVSPSPLEELIINHIVPEDQRONGSPFGWAFEGWVAALRSGLALRGL
TCLEASNHFPLPRDLIAQLPSRLYLDRNNSLVSLTYASFRNLTHLSLHLEDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYMDNVAWLKETEVPVDPKARLTCAPEK
MNRNGLDLNSDLDCCDAVLPSQTSYVFLGIVLALIGAILFLVLYLNRRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"
misc_feature
642. .1262
/genes="Tpbg"
/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [function unknown]"
/db_xref="CDD:COG4886"
1299. .1415
/genes="Tpbg"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 16.2 Length: 2423
Score: 44.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.7% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-12 (1-9) x BC058198 (1-2423)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||TCTTACCGGACACGATG 710
DB 684 AACCTTTCTTACCGGACACGATG 710

RESULT 31
AX961912
LOCUS AX961912 2557 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 123 from Patent WO03104277.
ACCESSION AX961912
VERSION AX961912.1 GI:40881322
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1
Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
Stat6 activation gene
Patent: WO 03104277-A 123 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
Location/Qualifiers
1. .2557
```

```
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556. .1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881323"
/translation="MPGAGSRGSPAGDGRRLRLRLALVLLGWVSASAPSSVPSSSTS
PAFLASGSAQPPPAERCPAAECSEAAATKVCNRRNLLVFPADLPVYVNRNLTGNQ
MTVLPAGAFARQPPPLADLEALNLSGNHLKEVCAGAFELPGLRLDLSHNPLTNSAF
VFAGSNASVSPSPLEELIINHIVPEDQRONGSPFGWAFEGWVAALRSGLALRGL
TCLEASNHFPLPRDLIAQLPSRLYLDRNNSLVSLTYASFRNLTHLSLHLEDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYMDNVAWLKETEVPVDPKARLTCAPEK
MNRNGLDLNSDLDCCDAVLPSQTSYVFLGIVLALIGAILFLVLYLNRRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 17.1 Length: 2557
Score: 44.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.7% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x AX961912 (1-2557)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
|||||TCTTACCGGACACGATG 864
DB 838 AACCTTTCTTACCGGACACGATG 864

RESULT 32
AX961914
LOCUS AX961914 2557 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 125 from Patent WO03104277.
ACCESSION AX961914
VERSION AX961914.1 GI:40881324
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1
Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
Stat6 activation gene
Patent: WO 03104277-A 125 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
Location/Qualifiers
1. .2557
```

CDS

```
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556. .1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881323"
/translation="MPGAGSRGSPAGDGRRLRLRLALVLLGWVSASAPSSVPSSSTS
PAFLASGSAQPPPAERCPAAECSEAAATKVCNRRNLLVFPADLPVYVNRNLTGNQ
MTVLPAGAFARQPPPLADLEALNLSGNHLKEVCAGAFELPGLRLDLSHNPLTNSAF
VFAGSNASVSPSPLEELIINHIVPEDQRONGSPFGWAFEGWVAALRSGLALRGL
TCLEASNHFPLPRDLIAQLPSRLYLDRNNSLVSLTYASFRNLTHLSLHLEDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYMDNVAWLKETEVPVDPKARLTCAPEK
MNRNGLDLNSDLDCCDAVLPSQTSYVFLGIVLALIGAILFLVLYLNRRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 17.1 Length: 2557
Score: 44.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
```

CDS

```
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556. .1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881325"
/translation="MPGAGSRGSPAGDGRRLRLRLALVLLGWVSASAPSSVPSSSTS
PAFLASGSAQPPPAERCPAAECSEAAATKVCNRRNLLVFPADLPVYVNRNLTGNQ
MTVLPAGAFARQPPPLADLEALNLSGNHLKEVCAGAFELPGLRLDLSHNPLTNSAF
VFAGSNASVSPSPLEELIINHIVPEDQRONGSPFGWAFEGWVAALRSGLALRGL
TCLEASNHFPLPRDLIAQLPSRLYLDRNNSLVSLTYASFRNLTHLSLHLEDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYMDNVAWLKETEVPVDPKARLTCAPEK
MNRNGLDLNSDLDCCDAVLPSQTSYVFLGIVLALIGAILFLVLYLNRRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 17.1 Length: 2557
Score: 44.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
```

```

Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.7% Indels: 0
DB: Gaps: 0

US-10-774-176-12 (1-9) x AX961914 (1-2557)

Qy 1 AsnLeuPheLeuThrGlyAenGlnLeu 9
Db 838 AACCTTTTCCTTACCGGCAACGATG 864

RESULT 33
MMU012160 7942 bp DNA linear ROD 15-APR-2005
DEFINITION Mus musculus 574 oncofetal trophoblast glycoprotein gene.
ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 574 gene; 574 oncofetal trophoblast glycoprotein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE Organisation of the mouse and human 574 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10365710
REFERENCE 2 (bases 1 to 7942)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK

FEATURES
source
1. .7942
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="Lambda Dash"
3108..3113
/bound_moiety="Sp1"
3114..3119
/bound_moiety="Sp1"
3124..5779
/gene="574"
3124..3151
/gene="574"
3152..3450
/gene="574"
3451..5779
/gene="574"
3779..5059
/gene="574"
/codon_start=1
/product="574 oncofetal trophoblast glycoprotein"
/protein_id="CAA09931.1"
/db_xref="GI:3805949"
/db_xref="GOA:Q920L0"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProt/TREMBL:Q920L0"
/translation="MGAGSRGPGAGDGLRLRLALVLGHWVSASAPSSVSSSTS
PADFLASGAQPPAPRCPCPCSEARATKVCVNRNLLSPADLPYVNRNLFITGQ
MTVLGAGAFARQPPADLEALNLSGNHLKEVCAGAFHFLGLRLDLNPLNLSAF
VFAGSASVAPSPLELNIHTVPPEDQRQNGSPFGMVAFGMAALRSGLALRGL
TRLRLASNIFLFLPDLIAQLPSLRILDLNNSLVLTASFRNLTHLSLHLELNL
KVLHNSFLAEWQGLAHVKVFLDNNPWVDCVMADVAVLAKETEVVDPKARLTCAFPK

```

```

MMNRGLDINSDDLDCDAVLPSQSLQTSYVFLGIVLALGALFLVLVLRKGIKKWMH
NIRDACDHMEGYHYRIEINADPLNTLSSNDV"

sig_peptide 3779..3865
mat_peptide 3866..5056
polyA_signal 5713..5718
polyA_signal 5759..5764

ORIGIN
Alignment Scores:
Pred. No.: 52.8 Length: 7942
Score: 44.00 Matches: 8
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.7% Indels: 0
DB: Gaps: 0

US-10-774-176-12 (1-9) x MMU012160 (1-7942)

Qy 1 AsnLeuPheLeuThrGlyAenGlnLeu 9
Db 4061 AACCTTTTCCTTACCGGCAACGATG 4087

RESULT 34
AL139393/c
LOCUS AL139393
DEFINITION Human DNA sequence from clone RP3-428L1.6 on chromosome 6q26-27
Contains part of a novel gene and the 5' end of the MAP3K4 gene for
mitogen-activated protein kinase kinase 4 (MTK1, MEKK4,
MAPKKK4, KIAA0213, homolog of yeast SSK2/SSK22 MAP kinase kinase
kinase). Contains a CpG island, complete sequence.
ACCESSION AL139393
VERSION AL139393.13 GI:11414475
KEYWORDS HTG; CpG island; KIAA0213; MAP3K4; MAPKKK4; MEKK4;
mitogen-activated protein kinase; MTK1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 85624)
Pelam, S.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Nov 28, 2000 this sequence version replaced gi:11228544.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

```


AL591045.5:9626..9704,AL591045.5:12933..13029,
 AL591045.5:17745..17845,AL591045.5:18552..18632,
 AL591045.5:22367..22558,AL591045.5:23752..23822,
 AL591045.5:24596..24712,AL591045.5:25608..25767,
 AL109942.13:2336..2458,AL109942.13:3168..3274,
 AL109942.13:7350..7370)
 /gene="MAP3K4"
 /locus_tag="RP3-473J16.4-002"
 /standard_name="OTHUMP0000017560"
 /notes="match: proteins: Q9P1M2"

Alignment Scores:
 Pred. No.: 562 Length: 85624
 Score: 44.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.7% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-12 (1-9) x AL139393 (1-85624)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 Db 49870 AACATTTTCTTAACGGCAATCAGTTG 49844

RESULT 35
 AC158516/c 167046 bp DNA linear ROD 21-JUN-2005
 LOCUS Mus musculus BAC clone RP24-511A23 from chromosome 9, complete
 DEFINITION
 AC158516 AC117768
 VERSION AC158516.2 GI:63025421
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 167046)
 Adams,S., Cotton,M. and Haglund,K.
 The sequence of Mus musculus BAC clone RP24-511A23
 Unpublished (2001)
 REFERENCE 2 (bases 1 to 167046)
 Wilson,R.K.
 Direct Submission
 Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 167046)
 Wilson,R.K.
 Direct Submission
 Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 167046)
 Wilson,R.K.
 Direct Submission
 Submitted (21-JUN-2005) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On May 4, 2005 this sequence version replaced gi:61656412.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: submissions@wustl.edu
 ----- Summary Statistics
 Center project name: M_BB0511A23
 Drafting center: WIBR

NOTICE:
 This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e. phred quality

>30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. The assembly was confirmed by restriction digest.
 This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal.
 If a sequence meets the criteria of the above statement, it needs no comments or tags. If the criteria are not met, such as ambiguous bases, then the region is duly annotated.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

SOURCE INFORMATION:
 The BAC Library has been constructed by Pieter de Jong and coworkers (http://www.chori.org) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at http://www.chori.org

This sequence is the entire insert of the clone.
 Location/Qualifiers
 1. 167046
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="9"
 /clone_lib="RP24-511A23"
 /clone="RP24-511A23"
 16685..16712
 /note="Sequence derived from PCR product of genomic DNA"
 31565..31779
 /note="Unresolved simple sequence repeat."
 46721..46808
 /note="Unresolved simple sequence repeat."
 142336..142347
 /note="Sequence derived from one plasmid subclone."

FEATURES

source
 1. 167046
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="9"
 /clone_lib="RP24-511A23"
 /clone="RP24-511A23"
 16685..16712
 /note="Sequence derived from PCR product of genomic DNA"
 31565..31779
 /note="Unresolved simple sequence repeat."
 46721..46808
 /note="Unresolved simple sequence repeat."
 142336..142347
 /note="Sequence derived from one plasmid subclone."

ORIGIN

Alignment Scores:
 Pred. No.: 1.09e+03 Length: 167046
 Score: 44.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.7% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-12 (1-9) x AC158516 (1-167046)

Qy 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9
 Db 110556 AACCTTTTCTTACCGCAACACGATG 110530

RESULT 36

AC015984/c 187098 bp DNA linear HTG 07-JUL-2000
 LOCUS AC015984
 DEFINITION Homo sapiens chromosome 6 clone RP11-570D18, WORKING DRAFT
 SEQUENCE, 14 unordered pieces.
 AC015984
 VERSION AC015984.5 GI:8568939
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE	Hominidae; Homo.
AUTHORS	1 (bases 1 to 187098)
TITLE	Waterston, R.H.
JOURNAL	The sequence of Homo sapiens clone
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 187098)
TITLE	Waterston, R.H.
JOURNAL	Direct Submission
COMMENT	Submitted (17-NOV-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA On Jun 16, 2000 this sequence version replaced gi:6955230.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0570D18
----- Summary Statistics -----
Sequencing vector: Mi3, 77k
Chemistry: Dye-terminator; plasmid; 23k
Chemistry: Dye-terminator; Big Dye; 23k of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 176778 bases at least Q40
Consensus quality: 180123 bases at least Q30
Consensus quality: 182248 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 185798; sum-of-contigs
Quality coverage: 5.05 in Q20 bases; agarose-fp
Quality coverage: 4.78 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	1	1313:	contig of 1313 bp in length
*	1314	1413:	gap of unknown length
*	1414	2545:	contig of 1132 bp in length
*	2546	2645:	gap of unknown length
*	2646	3998:	contig of 1353 bp in length
*	3999	4098:	gap of unknown length
*	4099	5141:	contig of 1043 bp in length
*	5142	5241:	gap of unknown length
*	5242	7495:	contig of 2254 bp in length
*	7496	7595:	gap of unknown length
*	7596	12363:	contig of 4768 bp in length
*	12364	12463:	gap of unknown length
*	12464	18289:	contig of 5826 bp in length
*	18290	18389:	gap of unknown length
*	18390	25557:	contig of 7178 bp in length
*	25568	25667:	gap of unknown length
*	25668	30467:	contig of 4800 bp in length
*	30468	30557:	gap of unknown length
*	30568	40583:	contig of 10016 bp in length
*	40584	40683:	gap of unknown length
*	40684	55115:	contig of 14432 bp in length
*	55116	55215:	gap of unknown length
*	55216	80201:	contig of 24986 bp in length
*	80202	80301:	gap of unknown length
*	80302	129362:	contig of 49061 bp in length
*	129363	129462:	gap of unknown length
*	129463	187098:	contig of 57636 bp in length.

FEATURES

	/chromosome="6"	
	/clone="RP11-570D18"	
	1. .1313	
misc_feature	/note="assembly_name:Contig4"	
gap	1314. .1413	
	/estimated_length=unknown	
misc_feature	1414. .2545	
gap	/note="assembly_name:Contig12"	
	2546. .2645	
	/estimated_length=unknown	
misc_feature	2646. .3998	
gap	/note="assembly_name:Contig13"	
	3999. .4098	
	/estimated_length=unknown	
misc_feature	4099. .5141	
gap	/note="assembly_name:Contig14"	
	5142. .5241	
	/estimated_length=unknown	
misc_feature	5242. .7495	
gap	/note="assembly_name:Contig15"	
	7496. .7595	
	/estimated_length=unknown	
misc_feature	7596. .12363	
gap	/note="assembly_name:Contig16"	
	12364. .12463	
	/estimated_length=unknown	
misc_feature	12464. .18289	
gap	/note="assembly_name:Contig17"	
	18290. .18389	
	/estimated_length=unknown	
misc_feature	18390. .25567	
gap	/note="assembly_name:Contig18"	
	25568. .25667	
	/estimated_length=unknown	
misc_feature	25668. .30467	
	/note="assembly_name:Contig19	
	clone_end:SP6	
gap	vector_side:right "	
	30468. .30567	
	/estimated_length=unknown	
misc_feature	30568. .40583	
gap	/note="assembly_name:Contig20"	
	40584. .40683	
	/estimated_length=unknown	
misc_feature	40684. .55115	
	/note="assembly_name:Contig21	
	clone_end:T7	
gap	vector_side:right "	
	55116. .55215	
	/estimated_length=unknown	
misc_feature	55216. .80201	
gap	/note="assembly_name:Contig22"	
	80202. .80301	
	/estimated_length=unknown	
misc_feature	80302. .129362	
gap	/note="assembly_name:Contig23"	
	129363. .129462	
	/estimated_length=unknown	
misc_feature	129463. .187098	
	/note="assembly_name:Contig24"	
ORIGIN		
Alignment Scores:		
Pred. No.:	1.22e+03	Length: 187098
Score:	44.00	Matches: 8
Percent Similarity:	100.0%	Conservative: 1
Best Local Similarity:	88.9%	Mismatches: 0
Query Match:	95.7%	Indels: 0
DB:	14	Gaps: 0
US-10-774-176-12 (1-9) x AC015984 (1-187098)		
Ov	1 AsnLeuPheLthrGlyAasnGlnLeu 9	

210237

/estimated_length=unknown

ORIGIN

Alignment Scores: 1.56e+03 239076
 Pred. No.: 44.00 Matches: 8
 Score: 44.00
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 95.7% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-12 (1-9) x AC106962 (1-239076)

QY 1 AsnLeuPheLeuThrGlyAsnGlnLeu 9

Db 16447 AACCTTTTCTCACTGGCAACGATG 16421
 |||||

RESULT 39

AY071256 1673 bp mRNA linear INV 20-DEC-2001
 LOCUS Drosophila melanogaster RE29447 full length cDNA.
 DEFINITION AY071256
 ACCESSION AY071256
 VERSION AY071256.1 GI:17945655
 KEYWORDS FLI_CDNA.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1673)
 Stapleton, M., Brokstein, P., Hong, L., Agbavani, A., Carlson, J.,
 Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,
 George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G.,
 Miranda, A., Mungall, C.J., Nuncio, J., Pacleb, J., Paragas, V., Park, S.,
 Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E., Rubin, G.M.
 and Celniker, S.
 Direct Submission
 Submitted (18-DEC-2001) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA

COMMENT

Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy, presence of a polyA tail and contiguity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcription of unspliced precursor RNAs, and
 reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our web site
 (<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES

source
 1..1673
 /location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="y; cn bw sp"
 /db_xref="taxon:7227"
 /map="100C1-100C2"
 1..1673
 /gene="CG15569"
 /note="alignment with genomic scaffold AB003778. gene does
 not completely overlap longest ORF"
 /db_xref="FLYBASE:FBgn0039845"
 45..1610
 /gene="CG15569"

gene

CDS

/note="Longest ORF"
 /codon_start=1
 /product="RE29447p"
 /protein_id="AAL48878.1"
 /db_xref="GI:17945656"
 /db_xref="FLYBASE:FBgn0039845"
 /translation="MEPIFLVCLLPQYSGARILAVFPLPSSSHYFALPYLKLAS
 LGHITSVSPQREPRNIDIPVPEVFENFNEVLRIASTPRSTWQSSDFINIVLNL
 LTKVLNNEGVRDILGPKQHPDLVIMDLWRMDVLVSLAAYFADPILIGMAYGTDMK
 IDBLMGNVSIYLSQSPSSFYDLEAYGERLLHLMERTFSYMYNKKRHRVKQRTLYSQ
 DFPVARKPLSISRNFDLVNQHFTLGPYPVPMQVGLVHDHSFEALSDEL
 DFIQAGSEGVLYESLGTWVKSLSSEDRKVLLETFAISLQRIVMKFEDELLPGKP
 PNVTISKNFFQQAIIAHPNVKLFIITHGGLLSTIESIHGKPMGLPCLFDQPRMDHV
 RYVGLGLVNLKIQMTSEPTIIRLTNKSFEETARITAAKYRDQPKMPEMTAIWMT
 EYVLSHKGAHQMVAGKDLGFVRYHSLDVFGLVGLVILGIVTYLLVMTLRKCLFL
 IKRGKCEAIKKIQ"

ORIGIN

Alignment Scores: 31.5 Length: 1673
 Pred. No.: 42.00 Matches: 8
 Score: 42.00
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.3% Indels: 0
 DB: 2 Gaps: 0

US-10-774-176-12 (1-9) x AY071256 (1-1673)

QY 1 AsnLeuPheLeuThrGlyAsnGln 8

Db 908 AATCTATTCTCACTGGGAACCAA 931
 |||||

RESULT 40

CQ611355 2820 bp DNA linear PAT 02-FEB-2004
 LOCUS CQ611355
 DEFINITION Sequence 39113 from Patent WO0171042.
 ACCESSION CQ611355
 VERSION CQ611355.1 GI:41662883
 KEYWORDS
 SOURCE Drosophila sp.
 ORGANISM Drosophila sp.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1
 Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
 Detection kits, such as nucleic acid arrays, for detecting the
 expression of 10,000 or more Drosophila genes and uses thereof
 Patent: WO 0171042-A 39113 27-SEP-2001;
 PE Corporation (NY) (US)

FEATURES

source
 1..2820
 /location/Qualifiers
 /organism="Drosophila sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:7242"

ORIGIN

Alignment Scores: 53 Length: 2820
 Pred. No.: 42.00 Matches: 8
 Score: 42.00
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.3% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x CQ611355 (1-2820)

QY 1 AsnLeuPheLeuThrGlyAsnGln 8

Db 864 AATCTATTCTCACTGGGAACCAA 887
 |||||

RESULT 41

CQ611354/c

LOCUS C0611354 5278 bp DNA linear PAT 02-FEB-2004
 DEFINITION Sequence 39112 from Patent WO0171042.
 ACCESSION C0611354
 VERSION C0611354.1 GI:41662882
 KEYWORDS
 SOURCE
 ORGANISM
 Drosophila sp.
 Drosophila sp.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
 1
 AUTHORS Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.
 TITLE Detection kits, such as nucleic acid arrays, for detecting the
 expression of 10,000 or more Drosophila genes and uses thereof
 JOURNAL Patent: WO 0171042-A 39112 27-SEP-2001;
 PE Corporation (NY) (US)
 FEATURES
 source
 1..5278
 Location/Qualifiers
 1..5278
 /organism="Drosophila sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:7242"

ORIGIN
 Alignment Scores:
 Pred. No.: 98.9 Length: 5278
 Score: 42.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.3% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-12 (1-9) x C0611354 (1-5278)

QY 1 AsnLeuPheLeuThrGlyAsnGln 8
 |||||
 3355 AATCTATTTCCTCTGGACCAA 3332

Db

RESULT 42
 AC101828/c
 LOCUS AC101828 68862 bp DNA linear HTG 23-NOV-2001
 DEFINITION Mus musculus clone RP24-428H10, LOW-PASS SEQUENCE SAMPLING.
 ACCESSION AC101828
 VERSION AC101828.1 GI:17060603
 KEYWORDS HTG; HTGS PHASE0.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.
 1 (bases 1 to 68862)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP24-428H10
 Unpublished
 2 (bases 1 to 68862)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galeagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Headford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kella, C., LaRocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McSwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
 Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L17490
 Center clone name: 428_H_10

 * NOTE: This record contains 88 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 670: contig of 670 bp in length
 * 671 770: gap of 100 bp
 * 771 1441: contig of 671 bp in length
 * 1442 1541: gap of 100 bp
 * 1542 2231: contig of 690 bp in length
 * 2232 2331: gap of 100 bp
 * 2332 3017: contig of 686 bp in length
 * 3018 3117: gap of 100 bp
 * 3118 3813: contig of 696 bp in length
 * 3814 3913: gap of 100 bp
 * 3914 4584: contig of 671 bp in length
 * 4585 4684: gap of 100 bp
 * 4685 5351: contig of 667 bp in length
 * 5352 5451: gap of 100 bp
 * 5452 6126: contig of 675 bp in length
 * 6127 6226: gap of 100 bp
 * 6227 6910: contig of 684 bp in length
 * 6911 7010: gap of 100 bp
 * 7011 7675: contig of 665 bp in length
 * 7676 7775: gap of 100 bp
 * 7776 8457: contig of 682 bp in length
 * 8458 8557: gap of 100 bp
 * 8558 9232: contig of 675 bp in length
 * 9233 9332: gap of 100 bp
 * 9333 10008: contig of 676 bp in length
 * 10009 10108: gap of 100 bp
 * 10109 10801: contig of 693 bp in length
 * 10802 10901: gap of 100 bp
 * 10902 11587: contig of 686 bp in length
 * 11588 11687: gap of 100 bp
 * 11688 12369: contig of 682 bp in length
 * 12370 12469: gap of 100 bp
 * 12470 13148: contig of 679 bp in length
 * 13149 13248: gap of 100 bp
 * 13249 13922: contig of 674 bp in length
 * 13923 14022: gap of 100 bp
 * 14023 14692: contig of 670 bp in length
 * 14693 14792: gap of 100 bp
 * 14793 15476: contig of 684 bp in length
 * 15477 15576: gap of 100 bp
 * 15577 16269: contig of 693 bp in length
 * 16270 16369: gap of 100 bp
 * 16370 17031: contig of 662 bp in length
 * 17032 17131: gap of 100 bp

TITLE
 JOURNAL

COMMENT

* 17132 17818: contig of 687 bp in length
* 17819 17918: gap of 100 bp
* 17919 18606: contig of 688 bp in length
* 18607 18706: gap of 100 bp
* 18707 19379: contig of 673 bp in length
* 19380 19479: gap of 100 bp
* 19480 20164: contig of 685 bp in length
* 20165 20264: gap of 100 bp
* 20265 20948: contig of 684 bp in length
* 20949 21048: gap of 100 bp
* 21049 21746: contig of 698 bp in length
* 21747 21846: gap of 100 bp
* 21847 22520: contig of 674 bp in length
* 22521 22620: gap of 100 bp
* 22621 23288: contig of 668 bp in length
* 23289 23388: gap of 100 bp
* 23389 24081: contig of 693 bp in length
* 24082 24181: gap of 100 bp
* 24182 24869: contig of 688 bp in length
* 24870 24969: gap of 100 bp
* 24970 25655: contig of 686 bp in length
* 25656 25755: gap of 100 bp
* 25756 26445: contig of 690 bp in length
* 26446 26545: gap of 100 bp
* 26546 27218: contig of 673 bp in length
* 27219 27318: gap of 100 bp
* 27319 28013: contig of 695 bp in length
* 28014 28113: gap of 100 bp
* 28114 28809: contig of 696 bp in length
* 28810 28909: gap of 100 bp
* 28910 29612: contig of 703 bp in length
* 29613 29712: gap of 100 bp
* 29713 30395: contig of 683 bp in length
* 30396 30495: gap of 100 bp
* 30496 31181: contig of 686 bp in length
* 31182 31281: gap of 100 bp
* 31282 31967: contig of 686 bp in length
* 31968 32067: gap of 100 bp
* 32068 32784: contig of 717 bp in length
* 32785 32884: gap of 100 bp
* 32885 33558: contig of 674 bp in length
* 33559 34335: gap of 100 bp
* 34336 34435: gap of 100 bp
* 34436 35126: contig of 691 bp in length
* 35127 35226: gap of 100 bp
* 35227 35915: contig of 689 bp in length
* 35916 36015: gap of 100 bp
* 36016 36701: contig of 686 bp in length
* 36702 36801: gap of 100 bp
* 36802 37507: contig of 706 bp in length
* 37508 37607: gap of 100 bp
* 37608 38297: contig of 690 bp in length
* 38298 38397: gap of 100 bp
* 38398 39089: contig of 692 bp in length
* 39090 39189: gap of 100 bp
* 39190 39854: contig of 665 bp in length
* 39855 39954: gap of 100 bp
* 39955 40647: contig of 693 bp in length
* 40648 40747: gap of 100 bp
* 40748 41432: contig of 685 bp in length
* 41433 41532: gap of 100 bp
* 41533 42228: contig of 696 bp in length
* 42229 42328: gap of 100 bp
* 42329 43007: contig of 679 bp in length
* 43008 43107: gap of 100 bp
* 43108 43793: contig of 686 bp in length
* 43794 43893: gap of 100 bp
* 43894 44514: contig of 621 bp in length
* 44515 44614: gap of 100 bp
* 44615 45274: contig of 660 bp in length
* 45275 45374: gap of 100 bp
* 45375 46054: contig of 680 bp in length

* 46055 46154: gap of 100 bp
* 46155 46843: contig of 689 bp in length
* 46844 46943: gap of 100 bp
* 46944 47633: contig of 690 bp in length
* 47634 47733: gap of 100 bp
* 47734 48420: contig of 687 bp in length
* 48421 48520: gap of 100 bp
* 48521 49211: contig of 691 bp in length
* 49212 49311: gap of 100 bp
* 49312 50016: contig of 705 bp in length
* 50017 50116: gap of 100 bp
* 50117 50824: contig of 708 bp in length
* 50825 50924: gap of 100 bp
* 50925 51600: contig of 676 bp in length
* 51601 51700: gap of 100 bp
* 51701 52384: contig of 684 bp in length
* 52385 52484: gap of 100 bp
* 52485 53161: contig of 677 bp in length
* 53162 53261: gap of 100 bp
* 53262 53934: contig of 673 bp in length

Alignment Scores:

Pred. No.: 1-27e+03 Length: 68862
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-12 (1-9) x AC101828 (1-68862)

Qy 1 AsnLeuPheLeuThrGlyAsnGln 8
| | | | | | | | | | | | | | | | | | | | | |
Db 65490 AATCTTTTTTACCGGCAACCAA 65467

RESULT 43

AC014787/c

LOCUS

AC014787 Drosophila melanogaster, 77707 bp DNA linear HTG 16-NOV-1999

DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.

AC014787

AC014787

AC014787.1 GI:6436548

HTG; HTGS_PHASE2.

KEYWORDS Drosophila melanogaster (fruit fly)

SOURCE Drosophila melanogaster

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 77707)

Adams, M. and Venter, J.C.

Direct Submission

Submitted (16-NOV-1999)

Rockville, MD, USA

This sequence was identified as CDM10212133 by the submitter.

For further information on this sequence e-mail to fly@celera.com.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

Location/Qualifiers

1..77707

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

ORIGIN

Alignment Scores:

Pred. No.: 1.43e+03 Length: 77707
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-12 (1-9) x AC014787 (1-77707)

QY 1 AsnLeuPheLeuThrGlyAsnGln 8
 Db 58770 AATCTATTTCCTACTGGGAACCA 58747

RESULT 44
 AC123802/c 121242 bp DNA linear ROD 05-NOV-2003
 LOCUS Mus musculus BAC clone RP24-398M15 from 3, complete sequence.
 DEFINITION AC123802
 ACCESSION AC123802
 VERSION AC123802.2 GI:22539238
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1 (bases 1 to 121242)
 Goyea, E., Meyer, R. and Schatzkamer, K.
 The sequence of Mus musculus BAC clone RP24-398M15
 Unpublished (2001)
 REFERENCE 2 (bases 1 to 121242)
 AUTHORS Wilson, R.
 TITLE Sequencing of Mus musculus
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 121242)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 121242)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 121242)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 REFERENCE 6 (bases 1 to 121242)
 AUTHORS Wilson, R.
 TITLE Direct Submission
 JOURNAL Submitted (05-NOV-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Aug 29, 2002 this sequence version replaced gi:21307495.

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics
 ----- Center project name: M_BB0398M15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see

<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES	source
1. 121242	Location/Qualifiers
/organism="Mus musculus"	
/mol_type="genomic DNA"	
/db_xref="taxon:10090"	
/chromosome="3"	
/map="3"	
/clone_lib="RPCI-24"	
/clone="RP24-398M15"	
12662..12817	repeat_region
/rpt_family="L1"	
15667..15796	repeat_region
/rpt_family="L1"	
16029..16238	repeat_region
/rpt_family="L1"	
18731..18918	repeat_region
/rpt_family="MIR"	
18943..19017	repeat_region
/rpt_family="ID"	
20445..20541	repeat_region
/rpt_family="MER1_type"	
22291..22449	repeat_region
/rpt_family="MER2_type"	
22445..22539	repeat_region
/rpt_family="MER2_type"	
24481..24963	repeat_region
/rpt_family="L1"	
26851..27088	repeat_region
/rpt_family="B4"	
27642..27731	repeat_region
/rpt_family="MER2_type"	
27782..27913	repeat_region
/rpt_family="L1"	
27917..28708	repeat_region
/rpt_family="L1"	
28887..28975	repeat_region
/rpt_family="Alu"	
33119..33690	repeat_region
/rpt_family="L1"	
33876..33937	repeat_region
/rpt_family="L1"	
33938..34227	repeat_region
/rpt_family="MaLR"	
34989..35547	repeat_region
/rpt_family="L1"	
35712..35762	repeat_region
/rpt_family="ERV1"	
35742..35788	repeat_region
/rpt_family="ERV1"	
35762..35812	repeat_region
/rpt_family="ERV1"	
36331..36579	repeat_region
/rpt_family="L1"	
38180..38274	repeat_region
/rpt_family="L1"	
40027..40124	repeat_region
/rpt_family="MER1_type"	
40970..41250	repeat_region
/rpt_family="B4"	
42225..42371	repeat_region
/rpt_family="L2"	
42964..43810	repeat_region
/rpt_family="L1"	

```

repeat_region 44038. .44142
/rpt_family="MaLR"
repeat_region 44363. .45315
/rpt_family="L1"
repeat_region 45592. .45907
/rpt_family="L1"
repeat_region 45977. .46345
/rpt_family="L1"
repeat_region 48623. .48752
/rpt_family="B4"
repeat_region 50213. .50292
/rpt_family="Alu"
repeat_region 55604. .55724
/rpt_family="L2"
repeat_region 55961. .57065
/rpt_family="L1"
repeat_region 57066. .57434
/rpt_family="MaLR"
repeat_region 57435. .57567
/rpt_family="MaLR"
repeat_region 57571. .58078
/rpt_family="L1"
repeat_region 58124. .58204
/rpt_family="L1"
repeat_region 58255. .58530
/rpt_family="L1"
repeat_region 58531. .58595
/rpt_family="Alu"
repeat_region 58596. .58654
/rpt_family="L1"
repeat_region 64834. .65023
/rpt_family="L1"
repeat_region 65048. .65181
/rpt_family="L1"
repeat_region 68250. .68394
/rpt_family="ERVL"
repeat_region 69754. .69930
/rpt_family="B2"
repeat_region 70221. .70360
/rpt_family="MaLR"
repeat_region 71033. .71419
/rpt_family="MaLR"
repeat_region 71582. .71718
/rpt_family="MaLR"
repeat_region 72136. .72409
/rpt_family="MaLR"
repeat_region 73269. .73415
/rpt_family="Alu"
repeat_region 74176. .74380
/rpt_family="B2"
repeat_region 74506. .74858
/rpt_family="L1"
repeat_region 76011. .76225
/rpt_family="B4"
repeat_region 76438. .76613
/rpt_family="L1"
repeat_region 76712. .76947
/rpt_family="B4"
repeat_region 77093. .77287
/rpt_family="L1"
repeat_region 78850. .78985
/rpt_family="L1"

```

```

Alignment Scores:
Pred. No.: 2,23e+03
Score: 42.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 91.3%
DB: 9

Length: 121242
Matches: 8
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-774-176-12 (1-9) x AC123802 (1-121242)

```

```

Qy 1 AsnLeuPheLeuThrGlyAsnGln 8
Db 103292 AATCTTTTATACCGCAACCAA 103269

RESULT 45
LOCUS AC083918
DEFINITION Homo sapiens chromosome 3 clone RP11-333B8 map 3, *** SEQUENCING IN
PROGRESS ***, 62 unordered pieces.
ACCESSION AC083918
VERSION AC083918.2 GI:12229341
KEYWORDS HTG; HTGS PHASE1.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 129302)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Cascie,A.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McSwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 15, 2001 this sequence version replaced gi:10697455.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11239
Center clone name: 333_B_8
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 74: contig of 74 bp in length
* 75 174: gap of 100 bp
* 175 915: contig of 741 bp in length
* 916 1015: gap of 100 bp
* 1016 1513: contig of 498 bp in length
* 1514 1613: gap of 100 bp
* 1614 2371: contig of 758 bp in length

```

* 2372	2471: gap of 100 bp	* 57444	60016: contig of 2573 bp in length
* 2472	3231: contig of 760 bp in length	* 60017	60116: gap of 100 bp
* 3232	3331: gap of 100 bp	* 60117	62464: contig of 2348 bp in length
* 3332	4455: contig of 1124 bp in length	* 62565	62564: gap of 100 bp
* 4456	4555: gap of 100 bp	* 62565	64478: contig of 1914 bp in length
* 4556	5438: contig of 883 bp in length	* 64479	64578: gap of 100 bp
* 5439	5538: gap of 100 bp	* 64579	66205: contig of 1627 bp in length
* 5539	6247: contig of 709 bp in length	* 66206	66305: gap of 100 bp
* 6248	6347: gap of 100 bp	* 66306	69316: contig of 3011 bp in length
* 6348	7868: contig of 1521 bp in length	* 69317	69416: gap of 100 bp
* 7869	7968: gap of 100 bp	* 69417	71575: contig of 2159 bp in length
* 7969	9424: contig of 1456 bp in length	* 71576	71675: gap of 100 bp
* 9425	9524: gap of 100 bp	* 71676	74218: contig of 2543 bp in length
* 9525	10365: contig of 841 bp in length	* 74219	74318: gap of 100 bp
* 10366	10465: gap of 100 bp	* 74319	76153: contig of 1835 bp in length
* 10466	11634: contig of 1169 bp in length	* 76154	76253: gap of 100 bp
* 11635	11734: gap of 100 bp	* 76254	78807: contig of 2554 bp in length
* 11735	12985: contig of 1251 bp in length	* 78808	78907: gap of 100 bp
* 12986	13085: gap of 100 bp	* 78908	81532: contig of 2625 bp in length
* 13086	14222: contig of 1137 bp in length	* 81533	81632: gap of 100 bp
* 14223	14322: gap of 100 bp	* 81633	84173: contig of 2541 bp in length
* 14323	15481: contig of 1159 bp in length	* 84174	84273: gap of 100 bp
* 15482	15581: gap of 100 bp	* 84274	87819: contig of 3546 bp in length
* 15582	16787: contig of 1306 bp in length	* 87820	87919: gap of 100 bp
* 16788	16887: gap of 100 bp	* 87920	91185: contig of 3266 bp in length
* 16888	17877: contig of 990 bp in length	* 91186	91285: gap of 100 bp
* 17878	17977: gap of 100 bp	* 91286	94191: contig of 2906 bp in length
* 17978	18936: contig of 959 bp in length	* 94191	94291: gap of 100 bp
* 18937	19036: gap of 100 bp	* 94292	98254: contig of 3963 bp in length
* 19037	20327: contig of 1291 bp in length	* 98255	98354: gap of 100 bp
* 20328	20427: gap of 100 bp	* 98355	101699: contig of 3345 bp in length
* 20428	21568: contig of 1141 bp in length	* 101700	101799: gap of 100 bp
* 21569	21668: gap of 100 bp	* 101800	105997: contig of 4198 bp in length
* 21669	23242: contig of 1574 bp in length	* 105998	106097: gap of 100 bp
* 23243	23342: gap of 100 bp	* 106098	110175: contig of 4078 bp in length
* 23343	24478: contig of 1136 bp in length	* 110176	110275: gap of 100 bp
* 24479	24578: gap of 100 bp	* 110276	115732: contig of 5457 bp in length
* 24579	26036: contig of 1458 bp in length	* 115733	115832: gap of 100 bp
* 26037	26136: gap of 100 bp	* 115833	120819: contig of 4987 bp in length
* 26137	27385: contig of 1249 bp in length	* 120820	120919: gap of 100 bp
* 27386	27485: gap of 100 bp	* 120920	125354: contig of 4435 bp in length
* 27486	28635: contig of 1150 bp in length	* 125355	125454: gap of 100 bp
* 28636	28735: gap of 100 bp	* 125455	129302: contig of 3848 bp in length.
* 28736	29870: contig of 1135 bp in length	FEATURES	
* 29871	31850: contig of 1880 bp in length	source	
* 31851	31950: gap of 100 bp	1. 129302	
* 31951	33650: contig of 1700 bp in length	/organism="Homo sapiens"	
* 33651	33750: gap of 100 bp	/mol_type="genomic DNA"	
* 33751	35262: contig of 1512 bp in length	/db_xref="taxon:9606"	
* 35263	35362: gap of 100 bp	/chromosome="3"	
* 35363	36898: contig of 1536 bp in length	/map="3"	
* 36899	36998: gap of 100 bp	/clone="RP11-333B8"	
* 36999	38634: contig of 1636 bp in length	/clone_lib="RPC1-11 Human Male BAC"	
* 38635	38734: gap of 100 bp	1. 74	
* 38735	40997: contig of 2263 bp in length	/note="assembly_fragment"	
* 40998	41097: gap of 100 bp	clone_end:SP6	
* 41098	42935: contig of 1838 bp in length	vector_side:left"	
* 42936	43035: gap of 100 bp	75. 174	
* 43036	44894: contig of 1859 bp in length	/estimated_length=100	
* 44895	44994: gap of 100 bp	175. 315	
* 44995	46507: contig of 1513 bp in length	/note="assembly_fragment"	
* 46508	46607: gap of 100 bp	Alignment Scores:	
* 46608	48188: contig of 1581 bp in length	Pred. No.: 2.38e+03 Length: 129302	
* 48189	48288: gap of 100 bp	Score: 42.00 Matches: 8	
* 48289	49922: contig of 1634 bp in length	Percent Similarity: 100.0% Conservative: 0	
* 49923	50022: gap of 100 bp	Best Local Similarity: 100.0% Mismatches: 0	
* 50023	52051: contig of 2029 bp in length	Query Match: 91.3% Indels: 0	
* 52052	52151: gap of 100 bp	DB: 14 Gaps: 0	
* 52152	54871: contig of 2720 bp in length	US-10-774-176-12 (1-9) x AC083918 (1-129302)	
* 54872	54971: gap of 100 bp	Qy 1 AsnLeuPheLeuThrGlyAsnGln 8	
* 54972	57343: contig of 2372 bp in length		
* 57344	57443: gap of 100 bp		

Db 31759 AACCTTTCTGACTGGAACCAA 31782

RESULT 46
AC021486
LOCUS
DEFINITION Homo sapiens chromosome 1, clone RP11-7316, complete sequence.
AC021486
VERSION AC021486.10 GI:17223358
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 144356)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A., Choepe, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 144356)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepe, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LeRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, T., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 1, 2001 this sequence version replaced gi:16948808.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>

FEATURES
source
Location/Qualifiers
1. 144356
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-7316"
/clone_lib="RPCI-11 Human Male BAC"
286..606
/rpt_family="AluSg"
repeat_region
655..678
/rpt_family="AT-rich"
complement(715..860)
/rpt_family="MIR3"
repeat_region
1676..1859
/rpt_family="MIR"
repeat_region
3382..3512
/rpt_family="MIR"
repeat_region
3532..3585
/rpt_family="AT-rich"
repeat_region
3632..3922
/rpt_family="AluSc"
complement(4398..4494)
/rpt_family="L2"
repeat_region
4495..4669
/rpt_family="AluSp"
4713..4736
/rpt_family="(TAAA)n"
complement(5213..5333)
/rpt_family="L1MB4A"
repeat_region
5350..5438
/rpt_family="(CA)n"
complement(5493..5583)
/rpt_family="AluJ/FRAM"
complement(5586..5998)
/rpt_family="L1MB4A"
repeat_region
6677..6762
/rpt_family="MIR"
complement(7431..7730)
/rpt_family="AluY"
complement(7748..8468)
/rpt_family="L1MA8"
8828..9179
/rpt_family="L2"
complement(12422..12550)
/rpt_family="L2"
complement(13639..14081)
/rpt_family="L1MB8"
15463..15743
/rpt_family="(TG)n"
15964..16047
/rpt_family="L1MB8"
16982..17289
/rpt_family="AluSg"
complement(18385..18434)
/rpt_family="MIR"
18602..18638
/rpt_family="(TG)n"
18936..18981
/rpt_family="MERSA"
complement(20595..20733)
/rpt_family="MIR"
complement(21009..21196)
/rpt_family="MIR3"
21316..21339
/rpt_family="AT-rich"
complement(21522..22371)
/rpt_family="L1MD3"
22374..22418
repeat_region

CONTACT: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I5353
Center clone name: 73_I_6

```

repeat_region /rpt_family="(TTATA)n"
22452. 22480
repeat_region /rpt_family="AT_rich"
22481. 22646
repeat_region /rpt_family="(TA)n"
complement(22684. 22827)
repeat_region /rpt_family="L1MC3"
complement(22833. 22909)
repeat_region /rpt_family="MIR"
23159. 23426
repeat_region /rpt_family="AluSx"
23432. 23452
repeat_region /rpt_family="(TAAA)n"
24111. 24553
repeat_region /rpt_family="MLTIC"
25193. 25335
repeat_region /rpt_family="MIR"
25458. 25537
repeat_region /rpt_family="L2"
complement(25689. 25854)
repeat_region /rpt_family="MLTIF1"
complement(26283. 26392)
repeat_region /rpt_family="L2"
26910. 26934
repeat_region /rpt_family="AT_rich"
27091. 27113
repeat_region /rpt_family="AT_rich"
complement(27119. 27412)
repeat_region /rpt_family="L1MA9"
28338. 28618
repeat_region /rpt_family="AluSg"
complement(30233. 30486)
repeat_region /rpt_family="MSTB1"
32821. 33059
repeat_region /rpt_family="L2"
33057. 33305
repeat_region /rpt_family="L2"
33339. 33497
repeat_region /rpt_family="L2"
34617. 34853
repeat_region /rpt_family="MIR"
complement(35197. 35474)
repeat_region /rpt_family="L2"
35587. 35800
repeat_region /rpt_family="L2"
36619. 36650
repeat_region /rpt_family="AT_rich"
38382. 38410
repeat_region /rpt_family="AT_rich"
complement(39539. 39719)
repeat_region /rpt_family="MIR"
complement(39880. 40035)
repeat_region /rpt_family="MIR3"
complement(40233. 40535)
repeat_region /rpt_family="MER4B"
40536. 40841
repeat_region /rpt_family="AluSx"
complement(40842. 41141)

```

```

Alignment Scores:
Pred. No.: 2.65e+03 Length: 144356
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0
DB: 8 Gaps: 0

```

US-10-774-176-12 (1-9) x AC021486 (1-144356)

Qy 1 AsnLeupheLeuthrGlyAsnGln 8

Db 130327 AACCTTTCTGACTGGAAACCAA 130350

RESULT 47

AC015567

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC015567 161674 bp DNA linear HTG 13-MAY-2001
Homo sapiens chromosome 18 clone RP11-398L3 map 18, WORKING DRAFT
SEQUENCE, 22 unordered pieces.

AC015567
AC015567.5 GI:14029904

HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Source: Homo sapiens (human)

Organism: Homo sapiens

Reference: Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 161674)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 18, clone RP11-398L3

Unpublished

2 (bases 1 to 161674)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Becker, R., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferrelira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tefaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 13, 2001 this sequence version replaced gi:9864762.

All repeats were identified using RepeatMasker:

Smit, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1016

Center clone name: 398 L 3

----- Summary Statistics

Sequencing vector: M13; M77815; 95% of reads

Sequencing vector: Plasmid; n/a; 5% of reads

Chemistry: Dye-terminator-amersham; 6% of reads

Chemistry: Dye-terminator Big Dye; 94% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 150819 bases at least Q40

Consensus quality: 155516 bases at least Q30

Consensus quality: 157646 bases at least Q20

Insert size: 170000; agarose-gel

Quality coverage: 159574; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 22 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 10093: contig of 10093 bp in length

* 10094 10193: gap of 100 bp

* 10194 11784: contig of 1591 bp in length

* 11785 11884: gap of 100 bp

* 11885 13970: contig of 2086 bp in length

* 13971 14070: gap of 100 bp

```

* 14071 16729: contig of 2659 bp in length
* 16730 16829: gap of 100 bp
* 16830 16920: contig of 2133 bp in length
* 16920 16962: contig of 2133 bp in length
* 16962 19062: gap of 100 bp
* 19062 21512: contig of 2450 bp in length
* 21512 21612: gap of 100 bp
* 21612 24514: contig of 2902 bp in length
* 24514 24614: gap of 100 bp
* 24614 28407: contig of 3793 bp in length
* 28407 28507: gap of 100 bp
* 28507 32001: contig of 3493 bp in length
* 32001 32101: gap of 100 bp
* 32101 36642: contig of 4542 bp in length
* 36642 36742: gap of 100 bp
* 36742 40704: contig of 3962 bp in length
* 40704 40804: gap of 100 bp
* 40804 45616: contig of 4812 bp in length
* 45616 45716: gap of 100 bp
* 45716 53538: contig of 7822 bp in length
* 53538 53638: gap of 100 bp
* 53638 62619: contig of 8981 bp in length
* 62619 62719: gap of 100 bp
* 62719 82101: contig of 19382 bp in length
* 82101 82201: gap of 100 bp
* 82201 90216: contig of 8015 bp in length
* 90216 90316: gap of 100 bp
* 90316 100558: contig of 10242 bp in length
* 100558 100659: gap of 100 bp
* 100659 110755: contig of 10097 bp in length
* 110755 110855: gap of 100 bp
* 110855 124345: contig of 13489 bp in length
* 124345 124444: gap of 100 bp
* 124444 137814: contig of 13370 bp in length
* 137814 137915: gap of 100 bp
* 137915 153435: contig of 15520 bp in length
* 153435 161674: gap of 100 bp
* 161674 161674: contig of 8140 bp in length.

FEATURES
    Location/Qualifiers
        1..161674
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="18"
            /map="18"
            /clone="RP11-398L3"
            /clone_lib="RPC1-11 Human Male BAC"
        1..10093
            /note="assembly_fragment"
            clone_end:SP6
            vector_side:left
            10094..10193
                /estimated_length=100
            10194..11784
                /note="assembly_fragment"
            11785..11884
                /estimated_length=100
            11885..13970
                /note="assembly_fragment"
            13971..14070
                /estimated_length=100
            14071..16729
                /note="assembly_fragment"
            16730..16829
                /estimated_length=100
            16830..18962
                /note="assembly_fragment"
            18963..19062
                /estimated_length=100
            19063..21512
                /note="assembly_fragment"
            21513..21612
                /estimated_length=100
            21613..24514

```

```

        /note="assembly_fragment"
        24515..24614
            /estimated_length=100
        24615..28407
            /note="assembly_fragment"
        28408..28507
            /estimated_length=100
        28508..32000
            /note="assembly_fragment"
        32001..32100
            /estimated_length=100
        32101..36642
            /note="assembly_fragment"
        36643..36742
            /estimated_length=100
        36743..40704
            /note="assembly_fragment"
        40705..40804
            /estimated_length=100
        40805..45616
            /note="assembly_fragment"
        45617..45716
            /estimated_length=100
        45717..53538
            /note="assembly_fragment"
        53539..53638
            /estimated_length=100
        53639..62619
            /note="assembly_fragment"
        62620..62719
            /estimated_length=100
        62720..82101
            /note="assembly_fragment"
        82102..82201
            /estimated_length=100
        82202..90216
            /note="assembly_fragment"
        90217..90316
            /estimated_length=100
        90317..100558
            /note="assembly_fragment"
        100559..100658
            /estimated_length=100
        100659..110755
            /note="assembly_fragment"
        110756..110855
            /estimated_length=100
        110856..124344
            /note="assembly_fragment"
        124345..124444
            /estimated_length=100
        124445..137814
            /note="assembly_fragment"
        137815..137914
            /estimated_length=100
        137915..153434
            /estimated_length=100

Alignment Scores:
Pred. No.:      2.97e+03      Length:      161674
Score:          42.00        Matches:      8
Percent Similarity: 100.0%    Conservative: 0
Best Local Similarity: 100.0% Mismatches:      0
Query Match:     91.3%       Indels:      0
DB:              14         Gaps:      0

US-10-774-176-12 (1-9) x AC015567 (1-161674)

Qy      1 AsnLeuPheLeuThrGlyAsnGln 8
Db      131591 AACCTTTCTGACTGGAAACCAA 131614
RESULT 48
AC019239/c

```

AC019239	165591 bp	DNA	linear	PRI 02-APR-2000
Homo sapiens chromosome 18, clone RP11-346H17, complete sequence.				
AC019239				
AC019239.5	GI:13324808			
HTG.				
Source	Homo sapiens (human)			
Organism	Homo sapiens			
Definition	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eumalia; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
Accession	1 (bases 1 to 165591)			
Version	Birren, B., Linton, L., Nusbaum, C. and Lander, E.			
Keywords	Homo sapiens chromosome 18, clone RP11-346H17			
Source	Unpublished			
Organism	2 (bases 1 to 165591)			
Definition	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, P., Boguslavskiy, L., Bouckgalter, B., Brown, A., Burkett, G., Castle, A., Choepell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArrelano, K., Dewar, K., Domino, M., Doyle, M., Feneser, J., Ferrario, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gaidyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liou, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurt, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.			
Title	Direct Submission			
Accession	Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
Version	3 (bases 1 to 165591)			
Keywords	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepell, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gaidyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazars, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Willson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.			
Title	Direct Submission			
Accession	Submitted (02-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
Version	On Mar 14, 2001 this sequence version replaced gi:11120835.			
Keywords	All repeats were identified using RepeatMasker:			
Source	Smit, A.P.A. & Green, P. (1996-1997)			
Organism	http://ftp.genome.washington.edu/RM/RepeatMasker.html			
Definition	----- Genome Center			
Title	Center: Whitehead Institute/ MIT Center for Genome Research			
Accession	Center code: WIBR			
Version	Web site: http://www-seq.wi.mit.edu			
Keywords	Contact: sequence submissions@genome.wi.mit.edu			
Source	----- Project Information			
Definition	Center project name: L1005			
Title	Center clone name: 346_H17			

```

repeat_region complement(27842..27986)
/rpt_family="L2"
repeat_region complement(28798..29139)
/rpt_family="AluSx"
repeat_region complement(29422..29807)
/rpt_family="MLT1J2"
repeat_region 30836..31119
/rpt_family="AluJo"
repeat_region 32388..32686
/rpt_family="AluJo"
repeat_region 32916..32978
/rpt_family="MIR3"
repeat_region complement(33183..33559)
/rpt_family="THE1C"
repeat_region 33787..33807
/rpt_family="AT-rich"
repeat_region 33974..34020
/rpt_family="(GA)n"
repeat_region complement(34429..34529)
/rpt_family="MIR"
repeat_region 36415..36468
/rpt_family="MIR"
repeat_region 37203..37613
/rpt_family="MER65B"
repeat_region complement(37614..37876)
/rpt_family="AluJo"
repeat_region 37877..37984
/rpt_family="MER65B"
repeat_region complement(38231..38261)
/rpt_family="L2"
repeat_region complement(38262..38689)
/rpt_family="MST4"
repeat_region complement(38690..38705)
/rpt_family="L2"
repeat_region 38797..38961
/rpt_family="MIR"
repeat_region 39298..39345
/rpt_family="(TC)n"
repeat_region complement(39820..39945)
/rpt_family="MIR3"
repeat_region 40305..40429
/rpt_family="MIR"
repeat_region 41521..41601
/rpt_family="MIR"
repeat_region 41652..41708
/rpt_family="AT-rich"
repeat_region complement(41912..42015)
/rpt_family="HAL1"
repeat_region complement(42151..42512)
/rpt_family="MLT1A1"
repeat_region complement(42646..43090)
/rpt_family="MLT1C"
repeat_region complement(43204..43798)
/rpt_family="HAL1"

```

```

Alignment Scores:
Pred. No.: 3.04e-03 Length: 165591
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0
DB: 8 Gaps: 0

```

US-10-774-176-12 (1-9) x AC019239 (1-165591)

```

QY 1 AsnLeuPheLeuThrGlyAsnGln 8
|||||
DB 145330 AACCTTTTCCTGACTGGAAACCA 145307

```

```

RESULT 49
AC102290/c AC102290 linear
LOCUS Mus musculus chromosome 3, clone RP24-359J7, complete sequence.
DEFINITION

```

AC102290
VERSION
KEYWORDS
SOURCE
ORGANISM

AC102290.7 GI:53793859
HTG.
Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 175819)

Birren,B., Nusbaum,C. and Lander,E.

Mus musculus chromosome 3, clone RP24-359J7

Unpublished

2 (bases 1 to 175819)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McSwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 175819)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Bouckgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., MacDonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

Direct Submission

Submitted (31-AUG-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 175819)

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavsky,L., Bouckgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,

Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, F., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (05-OCT-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 5, 2004 this sequence version replaced gi:51699743.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L18305

Center clone name: 359_J_7

FEATURES

source

Location/Qualifiers
1..175819
/organism="Mus musculus"
/molecule="genomic DNA"
/db_xref="taxon:10090"
/chromosome="3"
/map="3"
/clone="RP24-359J7"
/clone_lib="RPC1-24 Male Mouse BAC"

repeat_region

152..181
/rpt_family="AT_rich"

repeat_region

632..831
/rpt_family="Lx6"

repeat_region

833..868
/rpt_family="Lx6"

repeat_region

complement(1909..2010)
/rpt_family="Lx2"

repeat_region

2011..2036
/rpt_family="Lx2"

repeat_region

complement(2037..2282)
/rpt_family="Lx2"

repeat_region

3288..3339
/rpt_family="Lx2"

repeat_region

4055..4075
/rpt_family="AT_rich"

repeat_region

4514..4582
/rpt_family="TCCCn"

repeat_region

4582..4640
/rpt_family="TTCn"

repeat_region

complement(4886..4979)
/rpt_family="MLTJ1"

repeat_region

6452..6493
/rpt_family="Lx4"

repeat_region

7049..7078
/rpt_family="AT_rich"

repeat_region

7314..7563
/rpt_family="L1M4"

repeat_region

7641..7684
/rpt_family="AT_rich"

repeat_region

complement(7685..9375)
/rpt_family="Lx4"

repeat_region

9375..10647
/rpt_family="Lx4"

repeat_region

complement(10648..11066)
/rpt_family="IAPLTR2_MM"

repeat_region
11067..11132
/rpt_family="Lx4"

repeat_region
11134..11155
/rpt_family="AT_rich"

repeat_region
13239..13331
/rpt_family="CT-rich"

repeat_region
13851..13878
/rpt_family="(A)n"

repeat_region
14053..14117
/rpt_family="(CA)n"

repeat_region
15172..15202
/rpt_family="AT_rich"

repeat_region
15205..15350
/rpt_family="Bl_MM"

repeat_region
complement(15383..15663)
/rpt_family="Lx4"

repeat_region
15709..15736
/rpt_family="AT_rich"

repeat_region
complement(15830..17252)
/rpt_family="Lx4"

repeat_region
17253..17298
/rpt_family="(TA)n"

repeat_region
complement(17299..17357)
/rpt_family="Lx4"

repeat_region
complement(17364..17505)
/rpt_family="Bl_MM"

repeat_region
17511..17587
/rpt_family="AT_rich"

repeat_region
complement(17751..17897)
/rpt_family="RSINE1"

repeat_region
19167..19197
/rpt_family="AT_rich"

repeat_region
complement(19421..19759)
/rpt_family="ORR1A3"

repeat_region
21128..21172
/rpt_family="(TG)n"

repeat_region
22412..22508
/rpt_family="MIR"

repeat_region
22599..23045
/rpt_family="Lx2"

repeat_region
complement(23167..23446)
/rpt_family="MER20B"

repeat_region
complement(23742..24113)
/rpt_family="MTD"

repeat_region
24119..24280
/rpt_family="ORR1D"

repeat_region
24688..24830
/rpt_family="ORR1D"

repeat_region
24965..25336
/rpt_family="Lx6"

repeat_region
25337..25360
/rpt_family="(A)n"

Alignment Scores:

Pred. No.: 3.23e+03 Length: 175819
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-12 (1-9) x AC102290 (1-175819)

Qy 1 AsnLeuPheLeuThrGlyAsnGln 8

Db 11358 AATCTTTTITACCGCACCAA 11335

RESULT 50

AC008287/c

LOCUS

DEFINITION

AC008287

AC008287

AC008287

177480 bp DNA linear INV 24-FEB-2001
Drosophila melanogaster, chromosome 3R, region 100C-100C, BAC clone
BACR02G16, complete sequence.
AC008287 AC009390 AC009747

```

AC008287.4 GI:12957610
HTG.
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 177480)
AUTHORS
Celtniker, S.E., Adams, M.D., Krommiller, B., Tyler, D., Wan, K.H.,
Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,
Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,
Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M.,
Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D.,
Fierliera, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A.,
Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J.,
Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattai, B., Moshrefi, A.,
McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,
Pacaleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B.,
Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F.,
Stapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M.,
Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
Sequencing of Drosophila chromosome 3R, region 100C-100C
Unpublished
2 (bases 1 to 177480)
TITLE
Direct Submission
JOURNAL
Submitted (02-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
REFERENCE
Laboratory, MS 64-121, Berkeley, CA 94720, USA
AUTHORS
On or before Feb 24, 2001 this sequence version replaced
gi:6806808, gi:6630504, gi:6630510.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (http://www.fruitfly.org/sequence/) or send email
to bdg@fruitfly.berkeley.edu.
FEATURES
source
1..177480
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="100C-100C"
/clone="BACR02G16 (D1009)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBAC3.6)"
ORIGIN
Alignment Scores:
Pred. No.: 3.26e+03 Length: 177480
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-12 (1-9) x AC008287 (1-177480)
QY 1 AsnLeuPheLeuThrGlyAsnGln 8

```

Db 47455 AATCTATTCTCACTGGGAACCA 47432

Search completed: April 25, 2006, 20:41:36
Job time : 3090.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds
(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-11

Perfect score: 46

Sequence: 1 NLTEVPTDL 9

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/abes/ABSSWEB_spool/US10774176/runat_24042006_165112_19185/app_query.fasta_1
-DB=N Geneseq -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes05p
-USER=US10774176 @CGN 1.1 3463 @runat_24042006_165112_19185 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 21:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	505	14	ACL56146 Human col
2	46	100.0	927	6	ABT07721 Breast ca
3	46	100.0	927	8	ABX76333 Lung canc
4	46	100.0	927	10	ADB80503 Ovarian c

5	46	100.0	927	11	ADN38723	Adn38723 Cancer/an
6	46	100.0	973	8	AD56198	Ad56198 Human LRR
7	46	100.0	1156	6	ABV99349	ABV99349 Human NOV
8	46	100.0	1263	3	AA27058	Aa27058 Human S74
9	46	100.0	1331	8	AA56199	AAd56199 Human LRR
10	46	100.0	2020	10	ADJ56299	Adj56299 Human cDN
11	46	100.0	2053	8	ACC51052	Acc51052 Human bla
12	46	100.0	2053	8	ABX76332	ABx76332 Lung canc
13	46	100.0	2053	8	AD56197	Ad56197 Human LRR
14	46	100.0	2053	8	AD56200	Ad56200 Human LRR
15	46	100.0	2053	11	ADN38721	Adn38721 Cancer/an
16	46	100.0	2053	12	ADL06473	Adl06473 Human tum
17	46	100.0	2053	12	ADN03961	Adn03961 Antipsori
18	46	100.0	2053	13	ADR25444	Adr25444 Breast ca
19	46	100.0	2053	13	ACN38510	Acn38510 Tumour-as
20	46	100.0	2053	13	ADV35098	Adv35098 Human cDN
21	46	100.0	2338	5	AAS87175	Aas87175 DNA encod
22	46	100.0	2359	4	AAK94253	Aak94253 Human ful
23	46	100.0	2359	12	ADL30831	Adl30831 Full leng
24	46	100.0	2361	4	AAK94254	Aak94254 Human ful
25	46	100.0	2361	12	ADL26162	Adl26162 Human cDN
26	46	100.0	2361	12	ADL30833	Adl30833 Full leng
27	41	89.1	343	8	ABX40605	Abx40605 Bovine ES
28	41	89.1	1260	6	ABK87175	Abk87175 cDNA enco
29	41	89.1	1260	10	ADB97513	Adb97513 Feline ST
30	41	89.1	1260	10	ADB97452	Adb97452 DNA encod
31	41	89.1	1263	4	AAF89736	Aaf89736 Nucleotid
32	41	89.1	1263	6	ABK87174	Abk87174 cDNA enco
33	41	89.1	110000	12	ADN46845_12	Continuation (13 o
34	41	89.1	110000	12	ADN47591_08	Continuation (9 of
35	41	89.1	110000	12	ADN46123_12	Continuation (13 o
36	41	89.1	110000	12	ADN47209_08	Continuation (9 of
37	41	89.1	110000	12	ADN46464_12	Continuation (13 o
38	41	89.1	110000	12	ADN47960_08	Continuation (9 of
39	40	87.0	792	6	ABQ42606	Abq42606 Oligonucl
40	87.0	792	6	ABQ42607	Abq42607 Oligonucl	
41	40	87.0	795	6	ABQ42584	Abq42584 Oligonucl
42	40	87.0	795	6	ABQ42585	Abq42585 Oligonucl
43	39	84.8	691	6	ABT08922	Abt08922 Phase-1 R
44	39	84.8	691	10	ADG30834	Adg30834 Liver tox
45	39	84.8	691	12	ADG45384	Adg45384 Liver inf
46	39	84.8	691	12	ADH22691	Adh22691 Partial D
47	39	84.8	691	13	ADR91053	Adr91053 Spleen ne
48	39	84.8	933	6	ABK79398	Abk79398 Bacillus
49	39	84.8	2977	6	ABK63515	Abk63515 Rat seque
50	39	84.8	2977	10	ADB58652	Adb58652 Toxicity-
51	39	84.8	2977	10	ADB53331	Adb53331 Primary r
52	39	84.8	2977	10	ABT42240	Abt42240 Toxicity
53	39	84.8	2977	12	ADP72704	Adp72704 Renal tox
54	39	84.8	2977	13	ADRI5023	Adri5023 Rat elect
55	37	80.4	1881	6	ABQ70011	Abq70011 Listeria
56	37	80.4	1881	6	ABQ68026	Abq68026 Listeria
57	37	80.4	2001	12	ADP04686	Adp04686 Sea squir
58	37	80.4	110000	6	ABA03041_20	Continuation (21 o
59	37	80.4	110000	6	ABA03041_21	Continuation (22 o
60	36	78.3	80	12	ADM95692	Adm95692 Rat antis
61	36	78.3	303	10	ABX89079	Abx89079 Corn ear-
62	36	78.3	353	3	ABL80559	AbL80559 Human ova
63	36	78.3	359	3	AAA97767	Aaa97767 Chromatiu
64	36	78.3	434	8	ABX48404	Abx48404 Bovine ES
65	36	78.3	444	9	ACH28511	Ach28511 Human adu
66	36	78.3	608	6	ABL62076	AbL62076 Colon ade
67	36	78.3	863	4	AAK66602	Aak66602 Human imm
68	36	78.3	1657	13	ADX55059	Adx55059 Plant ful
69	36	78.3	1716	12	ADI43009	Adi43009 Plant tra
70	36	78.3	1716	12	ADO03140	Ado03140 Soybean o
71	36	78.3	1854	4	ABL02559	AbL02559 Drosophil
72	36	78.3	1854	4	ABL02559	AbL02559 Drosophil
73	36	78.3	1927	4	AAH98456	Aah98456 Human EST
74	36	78.3	2069	6	ABN59790	Abn59790 Novel hum
75	36	78.3	2366	2	AAQ06634	Aaq06634 Clone enc
76	36	78.3	2396	6	AAAL51004	AaL51004 Human zin
77	36	78.3	3854	4	ABL02558	AbL02558 Drosophil
78	36	78.3	4569	4	ABL03332	AbL03332 Drosophil

224	34	73.9	462	4	ABA23655	AbA23655 Probe #21	297	34	73.9	4900	10	ADP71518	Adf71518 Human Sli
225	34	73.9	462	4	AAK27620	Aak27620 Human bon	298	34	73.9	4900	11	ADM42030	Adm42030 Human Sli
226	34	73.9	462	4	AAK02173	Aak02173 Human bra	299	34	73.9	4950	2	AAx89162	Aax89162 Human sli
227	34	73.9	462	4	ABs27190	AbS27190 Human liv	300	34	73.9	4950	2	AAx14979	Aax14979 Nucleic a
228	34	73.9	462	5	AAI02112	AAi02112 Probe #21	301	34	73.9	4950	12	ADK70388	Adk70388 Respirato
229	34	73.9	462	6	ABs02084	AbS02084 Human gen	302	34	73.9	4950	12	ADN05841	Adn05841 Antipsoxi
230	34	73.9	464	13	ADo81477	ADo81477 Plant ful	303	34	73.9	5040	14	AEa17907	Aea17907 Human p21
231	34	73.9	556	4	AAI16955	AAi16955 Probe #68	c 304	34	73.9	5143	2	AAZ20131	Aaz20131 Human CDK
232	34	73.9	556	4	AAI16955	AAi16955 Probe #68	c 305	34	73.9	5143	2	AAZ20131	Aaz20131 Human CDK
233	34	73.9	556	4	AAI140949	AAi140949 Probe #96	306	34	73.9	5233	3	AAx15110	Aax15110 Transcrip
234	34	73.9	556	4	ABa28962	ABa28962 Probe #74	307	34	73.9	5233	3	AAc63530	AAc63530 Silt-2 pr
235	34	73.9	556	4	AAK35233	AAK35233 Human bon	308	34	73.9	5513	3	AAc63529	AAc63529 Silt prot
236	34	73.9	556	4	AAK09344	AAK09344 Human bra	309	34	73.9	5583	3	AAZ61825	AAZ61825 cDNA enco
237	34	73.9	556	4	AAK34979	AAK34979 Human liv	310	34	73.9	5583	3	AAc99758	AAc99758 Skin cell
238	34	73.9	556	6	ABs09653	ABs09653 Human gen	311	34	73.9	5583	6	ABL34910	ABl34910 Rat cDNA
239	34	73.9	561	4	AAI18911	AAi18911 Probe #88	312	34	73.9	6075	2	AAQ27451	AAq27451 Type A hu
240	34	73.9	561	4	AAI18911	AAi18911 Probe #88	313	34	73.9	6240	12	ADQ59531	ADq59531 Human can
241	34	73.9	561	4	AAI44032	AAi44032 Probe #12	314	34	73.9	6264	14	ADZ60757	ADz60757 Rat Pdgfr
242	34	73.9	561	4	ABa31089	ABa31089 Probe #95	315	34	73.9	6378	6	ABK35520	ABk35520 Human end
243	34	73.9	561	4	AAK38131	AAK38131 Human bon	316	34	73.9	6378	6	ABK68519	ABk68519 Kidney ca
244	34	73.9	561	4	AAK32418	AAK32418 Human bra	317	34	73.9	6378	6	ABL62372	ABl62372 Colon ade
245	34	73.9	561	4	ABs37752	ABs37752 Human liv	318	34	73.9	6378	10	ADP90648	ADp90648 Human hep
246	34	73.9	561	6	ABs12156	ABs12156 Human gen	319	34	73.9	6378	11	ADI31868	ADI31868 Human cDN
247	34	73.9	596	6	ABQ24688	ABq24688 Oligonuc1	320	34	73.9	6378	12	ADQ17832	ADq17832 Human sof
248	34	73.9	596	6	ABQ24689	ABq24689 Oligonuc1	321	34	73.9	6378	13	ADS83935	AdS83935 Human lym
249	34	73.9	675	2	AAQ76163	AAq76163 Human Sli	322	34	73.9	6412	2	AAQ05989	AAq05989 TF4 CDNA
250	34	73.9	805	6	ABQ46004	ABq46004 Oligonuc1	323	34	73.9	6428	5	ADL63502	ADl63502 Human ova
251	34	73.9	805	6	ABQ46005	ABq46005 Oligonuc1	324	34	73.9	6459	8	ABZ79546	ABz79546 Radish RE
252	34	73.9	812	6	AAZ15612	AAz15612 Human gen	325	34	73.9	6483	6	ABL33994	ABl33994 Human imm
253	34	73.9	861	8	ACC79236	ACC79236 Human GYP	326	34	73.9	6488	4	ABL10302	ABl10302 Drosophil
254	34	73.9	945	10	ADI21460	ADI21460 Novel hum	327	34	73.9	6499	12	ADQ59528	ADq59528 Human can
255	34	73.9	1059	4	AAI86705	AAi86705 Human pol	328	34	73.9	6556	14	ADZ13909	ADz13909 Murine ca
256	34	73.9	1174	3	AAc39984	AAc39984 Arabidops	329	34	73.9	6576	14	ADZ61190	ADz61190 Murine Pd
257	34	73.9	1210	8	ADAc68820	ADAc68820 Arabidops	330	34	73.9	6618	12	ADM69021	ADm69021 Human PPG
258	34	73.9	1279	13	ADr97621	ADr97621 1300 bp p	331	34	73.9	6618	12	ADM69032	ADm69032 Human PPG
259	34	73.9	1279	13	ADr97621	ADr97621 1300 bp p	332	34	73.9	6621	12	ADM69015	ADm69015 Human PPG
260	34	73.9	1431	10	ABZ40463	ABz40463 N. gonorr	333	34	73.9	6624	12	ADM69017	ADm69017 Human PPG
261	34	73.9	1518	6	ABV99762	ABv99762 Human PFK	334	34	73.9	6633	10	ADG89389	ADg89389 Cancer de
262	34	73.9	1767	10	ABs57785	ABs57785 Truncated	335	34	73.9	6633	12	ADM69030	ADm69030 Human PPG
263	34	73.9	1860	4	AAH16537	AAh16537 Human cDN	336	34	73.9	6633	12	ADM69034	ADm69034 Human PPG
264	34	73.9	2101	13	ADX52365	ADx52365 Plant ful	337	34	73.9	6633	12	ADM69011	ADm69011 Human pla
265	34	73.9	2113	13	ADX32389	ADx32389 Plant ful	338	34	73.9	6633	12	ADM69013	ADm69013 Human PPG
266	34	73.9	2385	6	ABZ11329	ABz11329 Human pol	339	34	73.9	6633	12	ADM69036	ADm69036 Human PPG
267	34	73.9	2385	12	ADW43847	ADw43847 Novel hum	340	34	73.9	6633	13	ADQ38813	ADq38813 Human SNP
268	34	73.9	2457	12	ADW72288	ADw72288 Porcine T	341	34	73.9	6633	14	ADY59388	ADy59388 Human PPG
269	34	73.9	2463	12	ADW72284	ADw72284 Rat TLR9	342	34	73.9	6633	14	ADZ13914	ADz13914 Human can
270	34	73.9	2501	10	ADB53960	ADb53960 CDKN1A ge	343	34	73.9	6639	12	ADM69019	ADm69019 Human PPG
271	34	73.9	2501	10	ADB84021	ADb84021 5' regula	c 344	34	73.9	6703	13	ADS11236	ADs11236 Human the
272	34	73.9	3099	12	ADW72283	ADw72283 Rat TLR9	345	34	73.9	7067	12	ADQ22465	ADq22465 Human sof
273	34	73.9	3270	10	ABs57782	ABs57782 Human pla	c 346	34	73.9	7369	8	ACF62753	ACf62753 Human p21
274	34	73.9	3270	10	ABs57782	ABs57782 Human pla	c 347	34	73.9	7369	8	ABZ09879	ABz09879 Human 5'
275	34	73.9	3270	10	ABs57781	ABs57781 Human pla	c 348	34	73.9	9235	6	ABZ09879	ABz09879 Human 5'
276	34	73.9	3270	10	ABs57780	ABs57780 cDNA enco	c 349	34	73.9	9235	6	ABs98926	ABs98926 Enterococ
277	34	73.9	3270	10	ABs57783	ABs57783 Human pla	c 350	34	73.9	9829	8	AAI60324	AAi60324 Aphid lec
278	34	73.9	3270	14	ADW42434	ADw42434 Human psy	c 351	34	73.9	10907	12	ACF57521	ACf57521 Human p21
279	34	73.9	3270	14	ADV44174	ADv44174 Human psy	c 352	34	73.9	10907	12	ADJ62879	ADj62879 Human p21
280	34	73.9	3282	8	ACA39329	ACa39329 prokaryot	c 353	34	73.9	10907	14	AEA60839	AEa60839 Human p21
281	34	73.9	3318	12	ADL03693	ADl03693 DNA enco	c 354	34	73.9	11948	4	AAE55234	AAe55234 Human p21
282	34	73.9	3329	14	ADW44365	ADw44365 Pig Toll-	c 355	34	73.9	11953	4	AAK90588	AAk90588 Human dig
283	34	73.9	3352	12	ADW72287	ADw72287 Porcine T	c 356	34	73.9	13480	4	AAK90589	AAk90589 Human dig
284	34	73.9	3457	11	ADW02539	ADw02539 Human cDN	c 357	34	73.9	13521	4	ABL03944	ABl03944 Drosophil
285	34	73.9	3477	13	ADW49653	ADw49653 Insulin s	c 358	34	73.9	14920	13	ADV87706	ADv87706 Streptoco
286	34	73.9	3663	13	ADT47331	ADt47331 Bacterial	359	34	73.9	14920	13	ADV78959	ADv78959 Streptoco
287	34	73.9	3746	14	ADZ62331	ADz62331 Murine 33	c 360	34	73.9	15300	14	ADY72597	ADy72597 A. orient
288	34	73.9	3927	6	ABN67045	ABn67045 Streptoco	c 361	34	73.9	15505	4	ABL19268	ABl19268 Drosophil
289	34	73.9	3971	12	ADQ35683	ADq35683 Novel mou	362	34	73.9	19130	12	ADQ97292	ADq97292 Mouse can
290	34	73.9	3972	13	ADV83635	ADv83635 Streptoco	363	34	73.9	22617	6	ABV99761	ABv99761 Human PFK
291	34	73.9	4054	2	AAQ06869	AAq06869 Sequence	c 364	34	73.9	22770	12	ADO43628	ADo43628 Nucleotid
292	34	73.9	4056	12	ADK42721	ADk42721 HEV parti	365	34	73.9	22863	3	AAa81500	AAa81500 N. mening
293	34	73.9	4575	2	AAK76161	AAK76161 Human Sli	366	34	73.9	23064	9	ADA01479	ADa01479 Human pur
294	34	73.9	4758	10	ADB85321	ADb85321 Human cDN	367	34	73.9	32203	14	ADZ59521	ADz59521 Secondary
295	34	73.9	4900	4	AAZ29462	AAz29462 Human Sli	368	34	73.9	36568	6	ABK50980	ABk50980 Human sol
296	34	73.9	4900	9	ADB90783	ADb90783 Human Sli	c 369	34	73.9	41648	4	AAE59551	AAe59551 Propionib

c 370	34	73.9	41648	8	ACF64480	Acf64480 Propionib	443	33	71.7	50	10	ADH59195	Adh59195 Human sec
371	34	73.9	41765	4	AAK76675	Aak76675 Human imm	444	33	71.7	50	10	ADH59195	Adh59195 Human sec
372	34	73.9	41772	4	AAK76675	Aak76675 Human imm	445	33	71.7	50	10	ACA58960	Aca58960 Human PRO
373	34	73.9	59475	12	ADQ59530	Adq59530 Human can	446	33	71.7	50	10	ACA58357	Aca58357 Probe #14
374	34	73.9	65952	12	ADQ59527	Adq59527 Human can	447	33	71.7	50	10	ADJ26242	Adj26242 Human sec
375	34	73.9	66009	14	ADZ13908	Adz13908 Murine ca	448	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 376	34	73.9	72750	3	AAH81468	Aah81468 N. mening	449	33	71.7	50	12	ADJ26242	Adj26242 Human sec
377	34	73.9	73599	14	ADY72575	Ady72575 Polyene p	450	33	71.7	50	12	ADJ26242	Adj26242 Human sec
378	34	73.9	89210	14	ADZ13911	Adz13911 Human can	451	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 379	34	73.9	99588	11	ACN45034	Acn45034 Human gen	452	33	71.7	50	12	ADJ26242	Adj26242 Human sec
380	34	73.9	99629	4	AAJ28550	Aaj28550 Genomic f	453	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 381	34	73.9	100944	12	ADQ59368	Adq59368 Human can	454	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 382	34	73.9	103747	6	ABQ88139	Abq88139 Human ost	455	33	71.7	50	12	ADJ26242	Adj26242 Human sec
383	34	73.9	110000	3	AAH81490_02	Continuation (3 of	456	33	71.7	50	12	ADJ26242	Adj26242 Human sec
384	34	73.9	110000	3	AAH81490_09	Continuation (10 of	457	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 385	34	73.9	110000	6	ABN71527_10	Continuation (11 of	458	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 386	34	73.9	110000	12	ADQ97050_2	Continuation (12 of	459	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 387	34	73.9	110000	13	ADV81204_11	Continuation (12 of	460	33	71.7	50	12	ADJ26242	Adj26242 Human sec
388	34	73.9	127432	12	ADQ43653	Ado43653 Nucleotid	461	33	71.7	50	12	ADJ26242	Adj26242 Human sec
389	34	73.9	164772	10	ADL13904	Adl13904 Osteoarth	462	33	71.7	50	12	ADJ26242	Adj26242 Human sec
390	34	73.9	191150	12	ADM69029	Adm69029 Human pla	463	33	71.7	50	12	ADJ26242	Adj26242 Human sec
391	34	73.9	271990	10	ADM25213	Adm25213 Fertility	464	33	71.7	50	12	ADJ26242	Adj26242 Human sec
392	34	73.9	271990	12	ADM61228	Adm61228 Radish nu	465	33	71.7	50	12	ADJ26242	Adj26242 Human sec
393	34	73.9	319608	3	AAH51601	Aah51601 Human chr	466	33	71.7	50	12	ADJ26242	Adj26242 Human sec
394	34	73.9	319608	5	AAH51601	Aah51601 Human chr	467	33	71.7	50	12	ADJ26242	Adj26242 Human sec
395	34	73.9	349980	3	AAJ21544	Aaj21544 Neisseria	468	33	71.7	50	12	ADJ26242	Adj26242 Human sec
396	34	73.9	349980	3	AAJ21544	Aaj21544 Neisseria	469	33	71.7	50	12	ADJ26242	Adj26242 Human sec
c 397	33	71.7	45	12	ADL24361	Adl24361 N meningi	470	33	71.7	50	12	ADJ26242	Adj26242 Human sec
398	33	71.7	50	2	AAH52319	Aah52319 Probe use	471	33	71.7	50	12	ADJ26242	Adj26242 Human sec
399	33	71.7	50	3	AAA49741	Aaa49741 Human PRO	472	33	71.7	50	12	ADJ26242	Adj26242 Human sec
400	33	71.7	50	3	ADJ78339	Adj78339 Human PRO	473	33	71.7	50	12	ADJ26242	Adj26242 Human sec
401	33	71.7	50	4	AAJ72477	Aaj72477 Human PRO	474	33	71.7	50	12	ADJ26242	Adj26242 Human sec
402	33	71.7	50	4	AAJ72477	Aaj72477 Human PRO	475	33	71.7	50	12	ADJ26242	Adj26242 Human sec
403	33	71.7	50	8	ACA97424	Ac97424 Human PRO	476	33	71.7	50	12	ADJ26242	Adj26242 Human sec
404	33	71.7	50	8	ACA97424	Ac97424 Human PRO	477	33	71.7	50	12	ADJ26242	Adj26242 Human sec
405	33	71.7	50	8	ACH06844	Ach06844 Human sec	478	33	71.7	50	12	ADJ26242	Adj26242 Human sec
406	33	71.7	50	8	ACH06844	Ach06844 Human sec	479	33	71.7	50	12	ADJ26242	Adj26242 Human sec
407	33	71.7	50	8	ABX71512	Abx71512 Human sec	480	33	71.7	50	12	ADJ26242	Adj26242 Human sec
408	33	71.7	50	8	ABX96081	Abx96081 Human sec	481	33	71.7	50	12	ADJ26242	Adj26242 Human sec
409	33	71.7	50	8	ACA05402	Ac05402 Human sec	482	33	71.7	50	12	ADJ26242	Adj26242 Human sec
410	33	71.7	50	8	ACD20069	Ac20069 Human sec	483	33	71.7	50	12	ADJ26242	Adj26242 Human sec
411	33	71.7	50	8	ACA54872	Ac54872 Novel sec	484	33	71.7	50	12	ADJ26242	Adj26242 Human sec
412	33	71.7	50	9	ACD19707	Adc19707 Human sec	485	33	71.7	50	12	ADJ26242	Adj26242 Human sec
413	33	71.7	50	9	ADJ29284	Adj29284 Human sec	486	33	71.7	50	12	ADJ26242	Adj26242 Human sec
414	33	71.7	50	9	ADJ29284	Adj29284 Human sec	487	33	71.7	50	12	ADJ26242	Adj26242 Human sec
415	33	71.7	50	9	ACD66854	Ac66854 Human sec	488	33	71.7	50	12	ADJ26242	Adj26242 Human sec
416	33	71.7	50	9	ACD83015	Ac83015 Human PRO	489	33	71.7	50	12	ADJ26242	Adj26242 Human sec
417	33	71.7	50	9	ADJ16115	Adj16115 Human sec	490	33	71.7	50	12	ADJ26242	Adj26242 Human sec
418	33	71.7	50	9	ACD23193	Ac23193 Human PRO	491	33	71.7	50	12	ADJ26242	Adj26242 Human sec
419	33	71.7	50	9	ADJ16539	Adj16539 Human sec	492	33	71.7	50	12	ADJ26242	Adj26242 Human sec
420	33	71.7	50	9	ADJ12968	Adj12968 Human sec	493	33	71.7	50	12	ADJ26242	Adj26242 Human sec
421	33	71.7	50	9	ADA41836	Ada41836 Human sec	494	33	71.7	50	12	ADJ26242	Adj26242 Human sec
422	33	71.7	50	9	ADA41836	Ada41836 Human sec	495	33	71.7	50	12	ADJ26242	Adj26242 Human sec
423	33	71.7	50	9	ADA42686	Ada42686 Human sec	496	33	71.7	50	12	ADJ26242	Adj26242 Human sec
424	33	71.7	50	9	ACD23555	Ac23555 Human PRO	497	33	71.7	50	12	ADJ26242	Adj26242 Human sec
425	33	71.7	50	10	ADJ77605	Adj77605 Human sec	498	33	71.7	50	12	ADJ26242	Adj26242 Human sec
426	33	71.7	50	10	ADJ77605	Adj77605 Human sec	499	33	71.7	50	12	ADJ26242	Adj26242 Human sec
427	33	71.7	50	10	ADJ77605	Adj77605 Human sec	500	33	71.7	50	12	ADJ26242	Adj26242 Human sec
428	33	71.7	50	10	ADJ77605	Adj77605 Human sec	501	33	71.7	50	12	ADJ26242	Adj26242 Human sec
429	33	71.7	50	10	ADJ77605	Adj77605 Human sec	502	33	71.7	50	12	ADJ26242	Adj26242 Human sec
430	33	71.7	50	10	ADJ77605	Adj77605 Human sec	503	33	71.7	50	12	ADJ26242	Adj26242 Human sec
431	33	71.7	50	10	ADJ77605	Adj77605 Human sec	504	33	71.7	50	12	ADJ26242	Adj26242 Human sec
432	33	71.7	50	10	ADJ77605	Adj77605 Human sec	505	33	71.7	50	12	ADJ26242	Adj26242 Human sec
433	33	71.7	50	10	ADJ77605	Adj77605 Human sec	506	33	71.7	50	12	ADJ26242	Adj26242 Human sec
434	33	71.7	50	10	ADJ77605	Adj77605 Human sec	507	33	71.7	50	12	ADJ26242	Adj26242 Human sec
435	33	71.7	50	10	ADJ77605	Adj77605 Human sec	508	33	71.7	50	12	ADJ26242	Adj26242 Human sec
436	33	71.7	50	10	ADJ77605	Adj77605 Human sec	509	33	71.7	50	12	ADJ26242	Adj26242 Human sec
437	33	71.7	50	10	ADJ77605	Adj77605 Human sec	510	33	71.7	50	12	ADJ26242	Adj26242 Human sec
438	33	71.7	50	10	ADJ77605	Adj77605 Human sec	511	33	71.7	50	12	ADJ26242	Adj26242 Human sec
439	33	71.7	50	10	ADJ77605	Adj77605 Human sec	512	33	71.7	50	12	ADJ26242	Adj26242 Human sec
440	33	71.7	50	10	ADJ77605	Adj77605 Human sec	513	33	71.7	50	12	ADJ26242	Adj26242 Human sec
441	33	71.7	50	10	ADJ77605	Adj77605 Human sec	514	33	71.7	50	12	ADJ26242	Adj26242 Human sec
442	33	71.7	50	10	ADJ77605	Adj77605 Human sec	515	33	71.7	50	12	ADJ26242	Adj26242 Human sec

516	33	71.7	553	12	ACH70070	Ach70070 Human gen	589	33	71.7	1305	9	ADA47681	Ada47681 Human PRO
517	33	71.7	572	6	ABQ19307	Abq19307 Oligonucle	590	33	71.7	1305	9	ADA18131	Ada18131 Human sec
518	33	71.7	572	6	ABQ19306	Abq19306 Oligonucle	591	33	71.7	1305	9	ACD66847	AcD66847 Human CDN
519	33	71.7	573	13	ADR91603	Adr91603 Novel S.	592	33	71.7	1305	9	ADA67476	Ada67476 Human PRO
520	33	71.7	573	14	ARA55473	Aea55473 Streptoco	593	33	71.7	1305	9	ADB30483	AdB30483 cDNA enco
521	33	71.7	578	13	ADO84262	Ado84262 Plant ful	594	33	71.7	1305	9	ADA85779	Ada85779 Novel hum
522	33	71.7	596	6	ABQ28939	Abq28939 Oligonucle	595	33	71.7	1305	9	ADA96991	Ada96991 Human PRO
523	33	71.7	596	6	ABQ28938	Abq28938 Oligonucle	596	33	71.7	1305	9	ADA79295	Ada79295 Human PRO
524	33	71.7	597	4	AAH29557	Aah29557 Drosophil	597	33	71.7	1305	9	ADA87434	Ada87434 Novel hum
525	33	71.7	654	10	ACA56345	Aca56345 Pig signa	598	33	71.7	1305	9	ADB16636	AdB16636 Human PRO
526	33	71.7	654	12	ADI56141	Adi56141 Human pol	599	33	71.7	1305	9	ACD83008	AcD83008 Human PRO
527	33	71.7	666	9	ADA13537	Ada13537 Human pro	600	33	71.7	1305	9	ADA16106	Ada16106 Human sec
528	33	71.7	699	10	ADF50996	Adf50996 Human HN1	601	33	71.7	1305	9	ADA91728	Ada91728 Novel hum
529	33	71.7	711	9	ADA13535	Ada13535 Human pro	602	33	71.7	1305	9	ADB14791	AdB14791 Human PRO
530	33	71.7	755	4	AAH05439	Aah05439 Human CDN	603	33	71.7	1305	9	ADB18752	AdB18752 Novel hum
531	33	71.7	759	6	ABQ16671	Abq16671 Oligonucle	604	33	71.7	1305	9	ADA93967	Ada93967 Human PRO
532	33	71.7	759	6	ABQ16670	Abq16670 Oligonucle	605	33	71.7	1305	9	ADB19863	AdB19863 Novel hum
533	33	71.7	760	6	ABS61523	AbS61523 Prostate	606	33	71.7	1305	9	ADB13175	AdB13175 Human PRO
534	33	71.7	765	9	ADA13539	Ada13539 Human pro	607	33	71.7	1305	9	ACD98575	AcD98575 Novel hum
535	33	71.7	771	5	AAF94094	Aaf94094 Primer sp	608	33	71.7	1305	9	ADA74429	Ada74429 Human PRO
536	33	71.7	771	14	ADY63521	Ady63521 Human clo	609	33	71.7	1305	9	ADA42251	Ada42251 Human sec
537	33	71.7	828	6	ABK79431	Abk79431 Bacillus	610	33	71.7	1305	9	ADB24662	AdB24662 Human PRO
538	33	71.7	830	2	AAT39039	Aat39039 Proteinas	611	33	71.7	1305	9	ADA82186	Ada82186 Human PRO
539	33	71.7	830	8	ACA64711	Aca64711 N. mening	612	33	71.7	1305	9	ADA75149	Ada75149 Human PRO
540	33	71.7	830	10	ADP43315	Adf43315 N. mening	613	33	71.7	1305	9	ADA85227	Ada85227 Novel hum
541	33	71.7	830	14	AEA03009	Aea03009 Neisseria	614	33	71.7	1305	9	ADA84675	Ada84675 Novel hum
542	33	71.7	858	9	ADA13541	Ada13541 Human pro	615	33	71.7	1305	9	ACD23186	AcD23186 Human PRO
543	33	71.7	858	9	ADA13553	Ada13553 Human pro	616	33	71.7	1305	9	ADB29931	AdB29931 cDNA enco
544	33	71.7	879	8	ABZ42278	Abz42278 Streptoco	617	33	71.7	1305	9	ADA80459	Ada80459 Human PRO
545	33	71.7	889	6	ABL59493	AbL59493 EST relat	618	33	71.7	1305	9	ADA75701	Ada75701 Human PRO
546	33	71.7	899	12	ADQ24119	Adq24119 Human sof	619	33	71.7	1305	9	ADA46926	Ada46926 Human PRO
547	33	71.7	903	3	AAC47757	Aac47757 Zea mays	620	33	71.7	1305	9	ADB25222	AdB25222 Human PRO
548	33	71.7	993	9	ADA13557	Ada13557 Human pro	621	33	71.7	1305	9	ADA93398	Ada93398 Human PRO
549	33	71.7	1029	4	ABR09491	AbR09491 Human PI	622	33	71.7	1305	9	ADB26748	AdB26748 cDNA enco
550	33	71.7	1131	11	ABD06337	Abd06337 Pseudomon	623	33	71.7	1305	9	ADB31035	AdB31035 cDNA enco
551	33	71.7	1139	6	ABQ33300	Abq33300 Oligonucle	624	33	71.7	1305	9	ADA60963	Ada60963 Homo sapi
552	33	71.7	1139	6	ABQ33301	Abq33301 Oligonucle	625	33	71.7	1305	9	ADB24110	AdB24110 Human PRO
553	33	71.7	1266	6	ABL59491	AbL59491 Nucleotid	626	33	71.7	1305	9	ADA96439	Ada96439 Human PRO
554	33	71.7	1266	12	ADL14124	AdL14124 Novel hum	627	33	71.7	1305	9	ADA81011	Ada81011 Human PRO
555	33	71.7	1275	10	ADB52811	AdB52811 Primary r	628	33	71.7	1305	9	ADA95887	Ada95887 Human PRO
556	33	71.7	1280	8	ABX72248	Abx72248 Human NOV	629	33	71.7	1305	9	ADB26196	AdB26196 cDNA enco
557	33	71.7	1280	8	ABX72247	Abx72247 Human NOV	630	33	71.7	1305	9	ADB21681	AdB21681 Novel hum
558	33	71.7	1302	3	ADC78390	Adc78390 Human PRO	631	33	71.7	1305	9	ADA77460	Ada77460 Human PRO
559	33	71.7	1305	2	AAX52227	Aax52227 Protein P	632	33	71.7	1305	9	ADB18200	AdB18200 cDNA enco
560	33	71.7	1305	3	AAA49720	Aaa49720 Human PRO	633	33	71.7	1305	9	ADA86883	Ada86883 Novel hum
561	33	71.7	1305	4	AAF72385	Aaf72385 Human PRO	634	33	71.7	1305	9	ADA16530	Ada16530 Human sec
562	33	71.7	1305	4	AAS21395	Aae21395 Human cDN	635	33	71.7	1305	9	ADA12959	Ada12959 Human sec
563	33	71.7	1305	4	AAC97421	Aac97421 Human ang	636	33	71.7	1305	9	ADA41827	Ada41827 Human sec
564	33	71.7	1305	6	ABL88081	AbL88081 Human PRO	637	33	71.7	1305	9	ADA87986	Ada87986 Novel hum
565	33	71.7	1305	6	ABK14009	Abk14009 Human clo	638	33	71.7	1305	9	ADA46374	Ada46374 Novel hum
566	33	71.7	1305	6	ABL95570	AbL95570 Human ang	639	33	71.7	1305	9	ADA17174	Ada17174 Human sec
567	33	71.7	1305	8	ACA60057	Aca60057 Human cDN	640	33	71.7	1305	9	ADA42677	Ada42677 Human sec
568	33	71.7	1305	8	ACD07457	AcD07457 Novel hum	641	33	71.7	1305	9	ADB28404	AdB28404 cDNA enco
569	33	71.7	1305	8	ACA03754	Aca03754 cDNA enco	642	33	71.7	1305	9	ADB28956	AdB28956 cDNA enco
570	33	71.7	1305	8	ABX71505	Abx71505 Human sec	643	33	71.7	1305	9	ADA76908	Ada76908 Human PRO
571	33	71.7	1305	8	ACH06837	Ach06837 Human sec	644	33	71.7	1305	9	ADA88538	Ada88538 Novel hum
572	33	71.7	1305	8	ABX89292	Abx89292 DNA encod	645	33	71.7	1305	9	ADA97543	Ada97543 Human PRO
573	33	71.7	1305	8	ACD41946	AcD41946 Human sec	646	33	71.7	1305	9	ADB27300	AdB27300 cDNA enco
574	33	71.7	1305	8	ABX96074	Abx96074 Human sec	647	33	71.7	1305	9	ADB22233	AdB22233 Novel hum
575	33	71.7	1305	8	ACA05395	Aca05395 cDNA enco	648	33	71.7	1305	9	ACD23548	AcD23548 Human PRO
576	33	71.7	1305	8	ACD20062	AcD20062 Human sec	649	33	71.7	1305	9	ADA66924	Ada66924 Human PRO
577	33	71.7	1305	8	ACA04175	Aca04175 Human cDN	650	33	71.7	1305	9	ADB22785	AdB22785 Human PRO
578	33	71.7	1305	8	ACA54865	Aca54865 Novel hum	651	33	71.7	1305	9	ADB23558	AdB23558 Human PRO
579	33	71.7	1305	9	ACD19700	AcD19700 Human sec	652	33	71.7	1305	9	ADA92280	Ada92280 Novel hum
580	33	71.7	1305	9	ADA45822	Ada45822 Novel hum	653	33	71.7	1305	9	ADB15343	AdB15343 Human PRO
581	33	71.7	1305	9	ADA76253	Ada76253 Human PRO	654	33	71.7	1305	9	ADB38595	AdB38595 Novel hum
582	33	71.7	1305	9	ADB29275	AdB29275 Human sec	655	33	71.7	1305	9	ADB38043	AdB38043 Novel hum
583	33	71.7	1305	9	ADA18903	Ada18903 Human PRO	656	33	71.7	1305	9	ADB66515	AdB66515 Novel hum
584	33	71.7	1305	9	ADA61526	Ada61526 Homo sapi	657	33	71.7	1305	10	ADB89595	AdB89595 Human PRO
585	33	71.7	1305	9	ADB19311	AdB19311 Novel hum	658	33	71.7	1305	10	ADB90327	AdB90327 Human PRO
586	33	71.7	1305	9	ADB27852	AdB27852 cDNA enco	659	33	71.7	1305	10	ADB77596	AdB77596 Human sec
587	33	71.7	1305	9	ADA86331	Ada86331 Novel hum	660	33	71.7	1305	10	ADB39428	AdB39428 Novel hum
588	33	71.7	1305	9	ADB15895	AdB15895 Human PRO	661	33	71.7	1305	10	ADB74732	AdB74732 Human sec

662	33	71.7	1305	10	ADB47051	Novel hum	735	33	71.7	1305	10	ADE34703	Human sec
663	33	71.7	1305	10	ADB86658	Human PRO	736	33	71.7	1305	10	ADE18857	Human PRO
664	33	71.7	1305	10	ADB77263	Novel hum	737	33	71.7	1305	10	ADE43053	Human PRO
665	33	71.7	1305	10	ADB34420	Human PRO	738	33	71.7	1305	10	ADD95842	Human PRO
666	33	71.7	1305	10	ADB35524	Human PRO	739	33	71.7	1305	10	AD227278	CDNA enco
667	33	71.7	1305	10	ADB33868	Human PRO	740	33	71.7	1305	10	ADD78846	CDNA enco
668	33	71.7	1305	10	ADB34972	Human PRO	741	33	71.7	1305	10	AD32796	Novel hum
669	33	71.7	1305	10	ADB36076	Human PRO	742	33	71.7	1305	10	AD42488	Human PRO
670	33	71.7	1305	10	ADB46471	Novel hum	743	33	71.7	1305	10	ADD80504	CDNA enco
671	33	71.7	1305	10	ADB28378	Human sec	744	33	71.7	1305	10	ADD89532	Human PRO
672	33	71.7	1305	10	ADC39578	Human sec	745	33	71.7	1305	10	AD40816	Human PRO
673	33	71.7	1305	10	ADC40092	Human sec	746	33	71.7	1305	10	AD404615	Human PRO
674	33	71.7	1305	10	ADC18920	Human sec	747	33	71.7	1305	10	AD92744	Human PRO
675	33	71.7	1305	10	ADC34216	Human sec	748	33	71.7	1305	10	AD321453	Novel hum
676	33	71.7	1305	10	ADC29271	Human sec	749	33	71.7	1305	10	ADG23094	Novel hum
677	33	71.7	1305	10	ADC28802	Human sec	750	33	71.7	1305	10	ADP97429	Human PRO
678	33	71.7	1305	10	ADC40687	Human sec	751	33	71.7	1305	10	ADG80493	Human PRO
679	33	71.7	1305	10	ADC19344	Human sec	752	33	71.7	1305	10	ADG79941	Human PRO
680	33	71.7	1305	10	ADC33792	Human sec	753	33	71.7	1305	10	ADH59186	Human sec
681	33	71.7	1305	10	ADC12862	Human sec	754	33	71.7	1305	10	ADH55233	Novel hum
682	33	71.7	1305	10	ADC50344	Novel hum	755	33	71.7	1305	10	ADH55785	Novel hum
683	33	71.7	1305	10	ADC71891	Novel hum	756	33	71.7	1305	10	AD137965	Human sec
684	33	71.7	1305	10	ADC59870	Novel hum	757	33	71.7	1305	10	AD164004	Novel hum
685	33	71.7	1305	10	ADC52877	Novel hum	758	33	71.7	1305	10	AD163452	Novel hum
686	33	71.7	1305	10	ADC57231	Novel hum	759	33	71.7	1305	10	ADH81866	Novel hum
687	33	71.7	1305	10	ADC60422	Novel hum	760	33	71.7	1305	10	ADH81314	Novel hum
688	33	71.7	1305	10	ADC50897	Novel hum	761	33	71.7	1305	10	ACA58953	Human PRO
689	33	71.7	1305	10	ADC65424	Human PRO	762	33	71.7	1305	10	ACD24004	Novel hum
690	33	71.7	1305	10	ADC54522	Novel hum	763	33	71.7	1305	10	ACA58350	CDNA enco
691	33	71.7	1305	10	ADC53483	Novel hum	764	33	71.7	1305	10	ACA67145	CDNA enco
692	33	71.7	1305	10	ADC59006	Novel hum	765	33	71.7	1305	10	ADJ26233	Human sec
693	33	71.7	1305	10	ADC55884	Novel hum	766	33	71.7	1305	11	ADM82483	Novel hum
694	33	71.7	1305	10	ADC58454	Novel hum	767	33	71.7	1305	11	ADM82483	Novel hum
695	33	71.7	1305	10	ADC12314	Human sec	768	33	71.7	1305	11	ADN15882	Novel hum
696	33	71.7	1305	10	ADC03128	Novel hum	769	33	71.7	1305	11	ADN16511	Novel hum
697	33	71.7	1305	10	ADC90120	Novel hum	770	33	71.7	1305	11	ADN15330	Novel hum
698	33	71.7	1305	10	ADC69539	CDNA enco	771	33	71.7	1305	11	ADN14778	Novel hum
699	33	71.7	1305	10	ADC48428	Human PRO	772	33	71.7	1305	11	AD164953	Novel hum
700	33	71.7	1305	10	ADC09957	Human PRO	773	33	71.7	1305	12	ADC81040	Novel hum
701	33	71.7	1305	10	ADD04532	Novel hum	774	33	71.7	1305	12	AD79148	Human sec
702	33	71.7	1305	10	ADC80488	Novel hum	775	33	71.7	1305	12	ADD76488	Human PRO
703	33	71.7	1305	10	ADD10995	Human PRO	776	33	71.7	1305	12	ADD87852	Human PRO
704	33	71.7	1305	10	ADD10308	Human sec	777	33	71.7	1305	12	ADD86256	Human PRO
705	33	71.7	1305	10	ADC47876	Human PRO	778	33	71.7	1305	12	AD79572	Human sec
706	33	71.7	1305	10	ADC04869	Human sec	779	33	71.7	1305	12	AD75704	Human PRO
707	33	71.7	1305	10	ADC79936	Novel hum	780	33	71.7	1305	12	AD73248	Human sec
708	33	71.7	1305	10	ADD11268	Human sec	781	33	71.7	1305	12	AD73248	Human sec
709	33	71.7	1305	10	ADD09405	Human PRO	782	33	71.7	1305	12	AD73248	Human sec
710	33	71.7	1305	10	ADD03875	Human sec	783	33	71.7	1305	12	AD73248	Human sec
711	33	71.7	1305	10	ADD03451	Human sec	784	33	71.7	1305	12	AD73248	Human sec
712	33	71.7	1305	10	ADD41118	Novel hum	785	33	71.7	1305	12	AD73248	Human sec
713	33	71.7	1305	10	ADD52257	CDNA enco	786	33	71.7	1305	12	AD73248	Human sec
714	33	71.7	1305	10	ADD52997	CDNA enco	787	33	71.7	1305	12	AD73248	Human sec
715	33	71.7	1305	10	ADD53549	Novel hum	788	33	71.7	1305	12	AD73248	Human sec
716	33	71.7	1305	10	ADD37061	Human sec	789	33	71.7	1305	12	AD73248	Human sec
717	33	71.7	1305	10	ADD51705	CDNA enco	790	33	71.7	1305	12	AD73248	Human sec
718	33	71.7	1305	10	ADD02504	Human PRO	791	33	71.7	1305	12	AD73248	Human sec
719	33	71.7	1305	10	ADD01938	Human PRO	792	33	71.7	1305	12	AD73248	Human sec
720	33	71.7	1305	10	ADD54120	Novel hum	793	33	71.7	1305	12	AD73248	Human sec
721	33	71.7	1305	10	ADD92437	Human PRO	794	33	71.7	1305	12	AD73248	Human sec
722	33	71.7	1305	10	ADD91333	Human PRO	795	33	71.7	1305	12	AD73248	Human sec
723	33	71.7	1305	10	ADD03947	Human PRO	796	33	71.7	1305	12	AD73248	Human sec
724	33	71.7	1305	10	AD32244	Novel hum	797	33	71.7	1305	12	AD73248	Human sec
725	33	71.7	1305	10	AD322176	CDNA enco	798	33	71.7	1305	12	AD73248	Human sec
726	33	71.7	1305	10	AD79400	CDNA enco	799	33	71.7	1305	12	AD73248	Human sec
727	33	71.7	1305	10	AD841936	Human PRO	800	33	71.7	1305	12	AD73248	Human sec
728	33	71.7	1305	10	AD817753	Human PRO	801	33	71.7	1305	12	AD73248	Human sec
729	33	71.7	1305	10	AD911885	Human PRO	802	33	71.7	1305	12	AD73248	Human sec
730	33	71.7	1305	10	AD833348	Novel hum	803	33	71.7	1305	12	AD73248	Human sec
731	33	71.7	1305	10	AD833900	Novel hum	804	33	71.7	1305	12	AD73248	Human sec
732	33	71.7	1305	10	AD79952	CDNA enco	805	33	71.7	1305	12	AD73248	Human sec
733	33	71.7	1305	10	ADD92989	Human PRO	806	33	71.7	1305	12	AD73248	Human sec
734	33	71.7	1305	10	AD819409	Human PRO	807	33	71.7	1305	12	AD73248	Human sec

PD 01-AUG-2002. OS Unidentified.
 XX
 PP 24-JAN-2002; 2002WO-US002242. FN WO200286443-A2.
 XX
 PR 24-JAN-2001; 2001US-0263965P. PD 31-OCT-2002.
 PR 02-FEB-2001; 2001US-0265928P. PP 18-APR-2002; 2002WO-US012476.
 PR 09-APR-2001; 2001US-00829472. PR 10-APR-2001; 2001US-0284770P.
 PR 09-APR-2001; 2001US-0282698P. PR 10-MAY-2001; 2001US-0290492P.
 PR 04-MAY-2001; 2001US-0288590P. PR 09-NOV-2001; 2001US-0339245P.
 PR 29-MAY-2001; 2001US-0294443P. PR 13-NOV-2001; 2001US-0350666P.
 XX (SOSB-) EOS BIOTECHNOLOGY INC. PR 29-NOV-2001; 2001US-0334370P.
 XX Mack DH, Gish KC, Afar D; PR 12-APR-2002; 2002US-0372246P.
 XX PI WPI; 2002-583738/62. PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX DR N-PSDB; ABJ05564. PI Aziz N, Murray R;
 XX PT Detecting a breast cancer-associated transcript in a patient's cell, PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.
 XX
 PS Claim 9; Page 372; 414pp; English.
 XX
 CC The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
 CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.21 Length: 927
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-11 (1-9) x ABT07721 (1-927)
 QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 241 AATCTGACCGAGGTGCCACGACCTG 267
 RESULT 3
 ABX76333
 ID ABX76333 standard; DNA; 927 BP.
 XX
 AC ABX76333;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polynucleotide #197.
 XX
 KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX

OS Unidentified.
 XX
 FN WO200286443-A2.
 XX
 PD 31-OCT-2002.
 XX
 PP 18-APR-2002; 2002WO-US012476.
 XX
 PR 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N, Murray R;
 XX
 DR WPI; 2003-093161/08.
 DR P-PSDB; ABU56604.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 CC Claim 22; Page 336; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.21 Length: 927
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-11 (1-9) x ABX76333 (1-927)
 QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 241 AATCTGACCGAGGTGCCACGACCTG 267
 RESULT 4
 ADB80503
 ID ADB80503 standard; DNA; 927 BP.
 XX
 AC ADB80503;
 XX
 DT 04-DEC-2003 (first entry)
 XX

DE Ovarian cancer-associated transcript #34.
 XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; ds, gene.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 1..927
 FT /*tag= a
 FT
 XX WO2002102235-A2.
 XX 27-DEC-2002.
 XX 18-JUN-2002; 2002WO-US019297.
 XX 18-JUN-2001; 2001US-0299234P.
 PR 27-AUG-2001; 2001US-0315287P.
 PR 05-SEP-2001; 2001US-0317544P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX Mack DH, Gish KC;
 XX WPI; 2003-167431/16.
 DR P-PSDB; ADB80504.
 XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.
 XX Claim 10; Page 297; 332pp; English.
 PS The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.
 XX Alignment Scores:
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Pred. No.: 4.21 Length: 927
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-11 (1-9) x ADB80503 (1-927)
 QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 241 MATCTGACCGAGGTGCCACGGACCTG 267
 RESULT 5
 ADN38723
 ID ADN38723 standard; cDNA; 927 BP.
 XX
 AC ADN38723;
 17-JUN-2004 (first entry)
 Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
 Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX Homo sapiens.
 OS WO2003042661-A2.
 XX 22-MAY-2003.
 PD 13-NOV-2002; 2002WO-US036810.
 XX 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-036809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX WPI; 2003-468649/44.
 DR P-PSDB; ADN38724.
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX Claim 8; SEQ ID NO 41; 1385pp; English.
 PS The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores: 4.21 Length: 927
Pred. No.: 46.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 11

US-10-774-176-11 (1-9) x ADN38723 (1-927)

Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 241 AATCTGACCGAGGTGCCACGACCTG 267

RESULT 6
AAD56198
ID AAD56198 standard; DNA; 973 BP.
XX
AC AAD56198;
XX
XX 07-AUG-2003 (first entry)
XX Human LRRCAPS related DNA #5.
XX Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
XX Homo sapiens.
XX
XX WO2003035831-A2.
XX
XX 01-MAY-2003.
XX
XX 21-OCT-2002; 2002WO-US033540.
XX
XX 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357600P.
PR 01-MAR-2002; 2002US-0361196P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
XX WPI; 2003-421410/39.
XX
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
XX Example 5; Page 74-75; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores: 4.45 Length: 973
Pred. No.: 46.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 8
DB: 8

US-10-774-176-11 (1-9) x AAD56198 (1-973)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 271 AATCTGACCGAGGTGCCACGACCTG 297

RESULT 7
ABV99349
ID ABV99349 standard; DNA; 1156 BP.
XX
AC ABV99349;
XX
XX 27-JAN-2003 (first entry)
XX Human NOV8a coding sequence.

XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antifertility; cerebroprotective; gene therapy; NOVx; NOV; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX Homo sapiens.

XX WO200272771-A2.
XX
XX 19-SEP-2002.
XX
XX 08-MAR-2002; 2002WO-US007288.
XX
XX 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277731P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 04-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.

```

PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 21-NOV-2001; 2001US-0333272P.
PR 03-DEC-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 04-DEC-2001; 2001US-0338092P.
PR 03-JAN-2002; 2001US-0337185P.
PR 08-MAR-2002; 2002US-00093463.
XX (CURA-) CURAGEN CORP.
XX
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
XX Boldog FI, Li L, Zerhusen BD, Tchernev VT, Gangelli EA, Verniet
XX Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
XX Voss EZ, Malyankar UM, Anderson DM, Patturajan M, Miller CE;
XX Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
XX Zhong M;
XX
XX WPI; 2002-732824/79.
XX P-P8DB; ABP70071.
XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,
XX diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
XX Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
XX disorders, and asthma.
XX
XX Claim 16; Page 114-115; 619pp; English.
XX
XX The present invention relates to new isolated proteins (NOVX) and their
XX coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
XX any number from 1 to 48. The NOVX proteins and coding sequences are
XX useful in the manufacture of a medicament for treating a syndrome
XX associated with a human disease, preferably a NOVX-associated disorder.
XX The NOVX coding sequences and proteins are useful for treating,
XX preventing or diagnosing diseases such as metabolic disorders, diabetes,
XX obesity, infectious disease, anorexia, cancer-associated cachexia,
XX cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
XX disease, immune disorders, hematopoietic disorders, cardiovascular
XX disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
XX disturbances associated with obesity, metabolic syndrome X or wasting
XX disorders associated with chronic diseases or various cancers. The NOVX
XX coding sequences and proteins may also be used as targets for the
XX identification of small molecules that modulate or inhibit e.g.
XX neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
XX wound healing and angiogenesis, in gene therapy, in generation of use
XX antibodies that bind immunospecifically to NOVX substances for use in
XX therapeutic or diagnostic methods
XX
XX SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.41 Length: 1156
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 241 AATCTGACCGAGTGCCACGACCTG 267

US-10-774-176-11 (1-9) x AAA27058 (1-1263)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 241 AATCTGACCGAGTGCCACGACCTG 267

Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x ABV99349 (1-1156)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 181 AATCTGACCGAGTGCCACGACCTG 207

RESULT 8
AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX
XX AAA27058;
XX
XX 22-AUG-2000 (first entry)
XX
XX Human 5T4 tumour-associated antigen gene.
XX
XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX ds.
XX
XX Homo sapiens.
XX
XX WO200029428-A2.
XX
XX 25-MAY-2000.
XX
XX 18-NOV-1999; 99WO-GB003859.
XX
XX 18-NOV-1998; 98GB-00025303.
XX
XX 27-JAN-1999; 99GB-00001739.
XX
XX 30-JUL-1999; 99GB-00017995.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll MW, Myers KA;
XX
XX WPI; 2000-387735/33.
XX
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX response useful in vaccinating against and in treating tumors.
XX
XX Example 2; Page 78; 79pp; English.
XX
XX The present sequence encodes the human 5T4 tumour-associated antigen
XX (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX carcinomas but has a highly restricted expression pattern in normal adult
XX tissues. It appears to be strongly correlated to metastasis in colorectal
XX and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX induced were inoculated with a virus expression vector containing the
XX present sequence. The 5T4 antigen was shown to be effective at eliciting
XX an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX the antigen and the antigen itself can be used to elicit an immune
XX response, preferably CTL or an antibody response in a subject
XX
XX SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.99 Length: 1263
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-11 (1-9) x AAA27058 (1-1263)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 241 AATCTGACCGAGTGCCACGACCTG 267

```

RESULT 9
 AAD56199
 ID AAD56199 standard; DNA; 1331 BP.
 XX AC
 XX AAD56199;
 XX 07-AUG-2003 (first entry)
 XX DT
 XX DE Human LRRCAPS related DNA #6.
 XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 XX KW LRRCAPS; cancer; gene therapy; da.
 XX OS Homo sapiens.
 XX PN WO2003035831-A2.
 XX PD 01-MAY-2003.
 XX PF 21-OCT-2002; 2002WO-US033540.
 XX PR 22-OCT-2001; 2001US-0338733P.
 XX PR 15-FEB-2002; 2002US-0357600P.
 XX PR 01-MAR-2002; 2002US-0361196P.
 XX PA (EXEL-) EXELIXIS INC.
 XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 XX PI Francis-Lang H, Friedman L;
 XX WPI; 2003-421410/39.
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX Disclosure; Page 75-76; 99pp; English.
 XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;
 SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.35 Length: 1331
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-11 (1-9) x AAD56199 (1-1331)
 QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 271 AATCTGACCGAGGTGCCACGACCTG 297
 RESULT 10
 ADJ56299
 ID ADJ56299 standard; cDNA; 2020 BP.
 XX AC
 XX ADJ56299;
 XX 06-MAY-2004 (first entry)
 XX DT

XX DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.
 XX KW human; differential expression; transactivator; proto-oncogene;
 KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
 KW MYCN activated cell.
 XX OS Homo sapiens.
 XX PN US2003119009-A1.
 XX PD 26-JUN-2003.
 XX PF 25-FEB-2002; 2002US-00084817.
 XX PR 23-FEB-2001; 2001US-0270784P.
 XX (STUA/) STUART S G.
 XX (NUCH/) NUCHTERN J G.
 XX (PLON/) PLON S E.
 XX (SHOH/) SHOHET J M.
 XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
 WPI; 2003-635698/60.
 XX New genes regulated by MYCN activation, useful in gene therapy,
 particularly for treating a subject with e.g. neuroblastoma or other
 cancers, or for diagnosing, staging or monitoring the treatment of the
 cancer.
 XX Claim 1; SEQ ID NO 105; 27pp; English.
 XX This invention relates to novel isolated cDNAs that are differentially
 CC expressed in MYCN activated cells. Specifically, it refers to
 CC polynucleotide sequences that exhibit differential expression patterns in
 CC cells activated by the transactivator MYCN, where MYCN is a proto-
 CC oncogene that is amplified in neuroblastoma cells and is common in small
 CC cell lung cancers. The present invention describes these cDNA molecules
 CC as useful for in hybridisation assays to detect expression of nucleic
 CC acids (or for complementary nucleic acids) in a present in a given sample, as
 CC well as for screening assays by identifying molecules or compounds that
 CC specifically bind the cDNA as a ligand and modulate function or activity.
 CC Accordingly, these compositions exhibit cytostatic activity and can also
 CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
 CC that is differentially expressed in MYCN activated cells, given in an
 CC exemplification of the invention. NOTE: This sequence does not appear in
 CC the printed specification but has been obtained in electronic format from
 CC the US Patent Office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.
 XX SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 10.2 Length: 2020
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-11 (1-9) x ADJ56299 (1-2020)
 QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 311 AATCTGACCGAGGTGCCACGACCTG 337
 RESULT 11
 ACC51052
 ID ACC51052 standard; cDNA; 2053 BP.
 XX AC
 XX ACC51052;
 XX DT

```

DT 12-JUN-2003 (first entry)
DE Human bladder cancer associated cDNA sequence SEQ ID NO:192.
KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
OS WO2003003906-A2.
PN 16-JAN-2003.
PD 03-JUL-2002; 2002WO-US021338.
XX 03-JUL-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX (BOSB-) EOS BIOTECHNOLOGY INC.
PA Mack DH, Aziz N;
XX
PI WPI; 2003-201532/19.
XX P-PSDB; ABR48236.
DR
DR
XX Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
XX Claim 6; Page 296; 307pp; English.
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridizes to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 10.4 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x ACC51052 (1-2053)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 325 AATCTGACCGAGGTGCCACGACCTG 351

RESULT 12
ABX76332
ID ABX76332 standard; DNA; 2053 BP.
XX
XX ABX76332;
XX
DT 02-APR-2003 (first entry)
XX

```

```

DE Lung cancer-associated polynucleotide #196.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
PN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
XX 18-APR-2002; 2002WO-US012476.
XX
XX 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
XX (BOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Aziz N, Murray R;
XX
XX WPI; 2003-093161/08.
DR P-PSDB; ABUS6603.
XX
XX Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
XX Claim 22; Page 335; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 10.4 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x ABX76332 (1-2053)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 325 AATCTGACCGAGGTGCCACGACCTG 351

```

```
RESULT 13
AAD56197
ID AAD56197 standard; DNA; 2053 BP.
XX
AC AAD56197;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human LRRCAPS DNA #11.
XX
KW Human, p53 pathway; Leucine rich repeat capricious related protein;
LRRCAPS; cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO2003035831-A2.
XX
PD 01-MAY-2003.
XX
PF 21-OCT-2002; 2002WO-US033540.
XX
PR 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357600P.
PR 01-MAR-2002; 2002US-0361196P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
DR WPI; 2003-421410/39.
XX
PT Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
PS Example 5; Page 73-74; 99pp; English.
XX
CC The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS DNA
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 10.4 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x AAD56197 (1-2053)
Oy 1 AsnLeuThrGluValProThrAspLeu 9
Db 325 AATCTGACCGAGGTGCCACGACCTG 351
RESULT 14
AAD56200
ID AAD56200 standard; DNA; 2053 BP.
XX
AC AAD56200;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human, differential expression; cancer; angiogenic disorder;
fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
inflammatory disease; autoimmune disease;
retinal neovascularisation syndrome; scarring; uterine fibroid;
detection; diagnosis; prognosis; drug screening; drug targeting;
wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX
```

```
DE Human LRRCAPS DNA #12.
XX
KW Human; p53 pathway; Leucine rich repeat capricious related protein;
LRRCAPS; cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO2003035831-A2.
XX
PD 01-MAY-2003.
XX
PF 21-OCT-2002; 2002WO-US033540.
XX
PR 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357600P.
PR 01-MAR-2002; 2002US-0361196P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
DR WPI; 2003-421410/39.
XX
PT Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
PS Disclosure; Page 76-77; 99pp; English.
XX
CC The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS DNA
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 10.4 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x AAD56200 (1-2053)
Oy 1 AsnLeuThrGluValProThrAspLeu 9
Db 325 AATCTGACCGAGGTGCCACGACCTG 351
RESULT 15
ADN38721
ID ADN38721 standard; cDNA; 2053 BP.
XX
AC ADN38721;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.
XX
KW Human; differential expression; cancer; angiogenic disorder;
fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
inflammatory disease; autoimmune disease;
retinal neovascularisation syndrome; scarring; uterine fibroid;
detection; diagnosis; prognosis; drug screening; drug targeting;
wound healing; contraception; cytostatic; cardiant; immunomodulatory;
```

KW vulnarary; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX PN WO2003042661-A2.

XX PD 22-MAY-2003.

XX PF 13-NOV-2002; 2002WO-US036810.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 21-NOV-2001; 2001US-0332464P.

XX PR 29-NOV-2001; 2001US-0334393P.

XX PR 03-DEC-2001; 2001US-0335394P.

XX PR 14-DEC-2001; 2001US-0340376P.

XX PR 08-JAN-2002; 2002US-0347211P.

XX PR 10-JAN-2002; 2002US-0347349P.

XX PR 08-FEB-2002; 2002US-0355250P.

XX PR 13-FEB-2002; 2002US-0356714P.

XX PR 20-FEB-2002; 2002US-0359077P.

XX PR 29-MAR-2002; 2002US-0368809P.

XX PR 04-APR-2002; 2002US-0370110P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PR 05-JUN-2002; 2002US-0386614P.

XX PR 16-JUL-2002; 2002US-0396839P.

XX PR 22-JUL-2002; 2002US-0397775P.

XX PR 22-JUL-2002; 2002US-0397845P.

XX PR 09-SEP-2002; 2002US-0409450P.

XX PA (BOBB-) EOS BIOTECHNOLOGY INC.

XX PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;

XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX DR WPI; 2003-468649/44.

XX DR P-PSDB; ADN38722.

XX PT Determining the presence or absence of a pathological cell in a patient,

XX PT useful for diagnosing, prognosing or treating cancer, comprises detecting

XX PT a nucleic acid in a biological sample.

XX PS Claim 8; SEQ ID NO 39; 1385pp; English.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	10.4	Length:	2053
Score:	46.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-774-176-11 (1-9) x ADN38721 (1-2053)

Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 325 AATCTGACCGAGGTGCCACGACCTG 351

RESULT 16

ADL06473

ID ADL06473 standard; cDNA; 2053 BP.

XX AC ADL06473;

XX XX 20-MAY-2004 (first entry)

XX DT Human tumour-associated antigenic target (TAT) cDNA sequence #53.

XX DE Human; tumour-associated antigenic target; TAT; cell death; tumour;

XX KW cancer; cytostatic; gene; ss.

XX OS Homo sapiens.

XX PN WO2004016225-A2.

XX XX 26-FEB-2004.

XX PF 19-AUG-2003; 2003WO-US025892.

XX PR 19-AUG-2002; 2002US-0404809P.

XX PR 21-AUG-2002; 2002US-0405645P.

XX PR 23-SEP-2002; 2002US-0413192P.

XX PR 15-OCT-2002; 2002US-0419008P.

XX PR 15-NOV-2002; 2002US-0426847P.

XX PR 02-JUL-2003; 2003US-0484959P.

XX PA (GETH) GENENTECH INC.

XX PI Desauvage PJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;

XX PI Spencer SD, Wu TD, Zhang Z;

XX DR WPI; 2004-257144/24.

XX DR P-PSDB; ADL06552.

XX PT New antibody that binds to a tumor-associated antigenic target (TAT)

XX PT polypeptide, useful for preparing a composition for diagnosing or

XX PT treating cancer.

XX PS Claim 1; SEQ ID NO 53; 319pp; English.

XX CC The present invention relates to the isolation of human tumour-associated

XX CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also

XX CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is

XX CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a

XX CC humanised antibody. It is conjugated to a growth inhibitory agent. It is

XX CC produced in bacteria or in CHO cells and induces death of a cell to which

XX CC it binds. The antibody is useful for preparing a composition for

XX CC diagnosing or treating tumours and cancer. The present sequence

XX CC represents a human TAT cDNA sequence of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	10.4	Length:	2053
Score:	46.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-11 (1-9) x ADL06473 (1-2053)

Qy 1 AsnLeuThrGluValProThrAspLeu 9

Db 325 AATCTGACCGAGGTGCCACGACCTG 351

RESULT 17


```
XX (GETH ) GENENTECH INC.
XX
XX Wu TD, Zhang Z, Zhou Y;
PI
XX WPI; 2004-347921/32.
DR P-PSDB; ABM80804.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
XX Claim 1; SEQ ID NO 2070; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acids and
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods and compositions for the treatment or
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX antibodies, antagonists, binding molecules and compositions are useful
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention
XX
XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 10.4 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0
US-10-774-176-11 (1-9) x ACN38510 (1-2053)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 325 AATCTGACCGAGGTGCCACGACCTG 351
RESULT 20
ADV35098
ID ADV35098 standard; cDNA; 2053 BP.
XX
XX ADV35098;
AC
XX
XX 10-FEB-2005 (first entry)
DT
XX
XX Human cDNA of an exemplary efficacy gene for BAD SeqID174.
DE
XX
XX human; ss; multi-parameter high throughput screening; MPHTS;
KW disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
KW bipolar affective disorder; BAD; autism; Parkinson's;
KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.
XX
XX Homo sapiens.
OS
XX
XX US2003096264-A1.
PN
XX
XX 22-MAY-2003.
PD
XX
```

```
PF 18-JUN-2002; 2002US-00175523.
XX
XX 18-JUN-2001; 2001US-0299151P.
PR
XX 07-SEP-2001; 2001US-0317828P.
PR
XX 25-SEP-2001; 2001US-0325150P.
PR
XX 14-NOV-2001; 2001US-0333047P.
PR
XX 18-JAN-2002; 2002US-0349936P.
PR
XX 04-MAR-2002; 2002US-0361834P.
XX
XX (PSYC-) PSYCHIATRIC GENOMICS INC.
PA
XX Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
XX Palfreyman M, Rajan P;
PI
XX WPI; 2004-118903/12.
XX
XX Identifying a compound that can treat disease or disorders, such as, a
XX neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
XX determining the expression of one or more efficacy genes in a cell
XX contacted with the test compound.
XX
XX Example 6; SEQ ID NO 174; 39pp; English.
XX
XX This invention relates to a novel screening method identified as a multi-
XX parameter high throughput screening (MPHTS) assay. Specifically, it
XX refers to an assay that utilises the disease signature of a plurality of
XX specific genes associated with a particular disease, and identifies
XX differential expression between those cells taken from individuals
XX affected by that disease and those that are not affected. The present
XX invention then describes the screening of candidate pharmaceutical
XX compounds to identify those that have a potential therapeutic benefit for
XX the treatment of neuropsychiatric and neurodegenerative disorders
XX including schizophrenia, bipolar affective disorder (BAD) and autism, as
XX well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
XX of this invention exhibit various activities including neuroleptic,
XX nootropic, antimanic and antidepressant. Furthermore, the screening
XX method used in MPHTS will be automated, such that a large number of test
XX compounds may be rapidly screened with a minimal amount of labour and
XX effort. This polynucleotide is a human cDNA sequence of a gene that is
XX differentially expressed in the presence of a therapeutic compound and
XX represents an exemplary efficacy gene for bipolar affective disorder,
XX given in an exemplification of the invention.
XX
XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 10.4 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0
US-10-774-176-11 (1-9) x ADV35098 (1-2053)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 325 AATCTGACCGAGGTGCCACGACCTG 351
RESULT 21
AAS87175
ID AAS87175 standard; cDNA; 2338 BP.
XX
XX AAS87175;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #22979.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
```

```

XX WO200175067-A2.
XX
XX
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG22988.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 22979; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 12.1 Length: 2338
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-11 (1-9) x AAS87175 (1-2338)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 582 AATCTGACCGAGGTGCCACGGACCTG 608

RESULT 22
AAK94253
ID AAK94253 standard; cDNA; 2359 BP.
XX
XX AC AAK94253;
XX
XX 06-NOV-2001 (first entry)
XX
XX DE Human full-length cDNA, SEQ ID NO: 2864.
XX
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX OS Homo sapiens.
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX
XX 11-JAN-2000; 2000JP-00118774.
XX
XX 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
XX
XX P-PSDB; AAM93333.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesising full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5' and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesising the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a full length human cDNA of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in CD-ROM format directly
XX from BPO
XX
XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 12.2 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-11 (1-9) x AAK94253 (1-2359)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 664 AATCTGACCGAGGTGCCACGGACCTG 690

RESULT 23
ADL30831
ID ADL30831 standard; cDNA; 2359 BP.
XX
XX AC ADL30831;
XX
XX 20-MAY-2004 (first entry)
XX
XX DE Full length human cDNA clone SeqID 2864.
XX
XX KW human; medicine; signal transduction; glycoprotein; transcription;
XX oligo-capping method; ss; gene.
XX
XX OS Homo sapiens.
XX
XX EP1396543-A2.
XX
XX 10-MAR-2004.

```

```

XX PF 07-JUL-2000; 2003EP-00025638.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183865.
XX PR 07-JUL-2000; 2000EP-00114089.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX DR WPI; 2004-204755/20.
XX DR P-PSDB; ADL30832.
XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
XX PT length human cDNAs.
XX PS Example 1; SEQ ID NO 2864; 1340pp; English.
XX CC This invention relates to a novel primers useful for synthesizing full
XX CC length cDNA molecules that encode human proteins. Specifically, it refers
XX CC to secretory or membrane proteins that are potential therapeutic agents/
XX CC target molecules in the field of medicine, and in particular genes
XX CC encoding proteins that are associated with signal transduction,
XX CC glycoproteins and transcription. The present invention describes a method
XX CC for efficiently cloning a full length human cDNA from both the 5' and 3'
XX CC ends using the oligo-capping method. This polynucleotide sequence is a
XX CC full length human cDNA clone of the invention.
XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 12.2 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADL30831 (1-2359)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 664 AATCTGACCGAGGTGCCCGACCTG 690

RESULT 24
AAK94254
ID AAK94254 standard; cDNA; 2361 BP.
XX AC AAK94254;
XX DT 06-NOV-2001 (first entry)
XX DE Human full-length cDNA, SEQ ID NO: 2866.
XX KW Human, full length cDNA; cDNA synthesis; oligo-capping; ss.
XX OS Homo sapiens.
XX PN EP130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX PA (HELI-) HELIX RES INST.
XX PA

PI Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX DR WPI; 2001-524255/58.
XX DR P-PSDB; AAM93334.
XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX PS Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.
XX CC The invention relates to primers for synthesizing full length cDNA
XX CC clones. 830 cDNA molecules encoding a human protein have been isolated
XX CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX CC been determined. Primers for synthesizing the full length cDNA are useful
XX CC for clarifying the function of the protein encoded by the cDNA. The full
XX CC length clones were obtained by construction of full length enriched cDNA
XX CC libraries that were synthesised by the oligo-capping method. The primers
XX CC enable the production of the full length cDNA easily without any special
XX CC methods. The present sequence is a full length human cDNA of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in CD-ROM format directly
XX CC from EPO
XX SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 12.2 Length: 2361
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-11 (1-9) x AAK94254 (1-2361)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 666 AATCTGACCGAGGTGCCCGACCTG 692

RESULT 25
ADI26162
ID ADI26162 standard; cDNA; 2361 BP.
XX AC ADI26162;
XX DT 22-APR-2004 (first entry)
XX DE Human cDNA encoding protein that promotes STAT6 activation #64.
XX KW ss; gene; human; signal transducer and activator of transcription 6;
XX KW STAT6; immunogen; STAT6 activation; allergy; inflammation; cancer;
XX KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
XX KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
XX KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
XX KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
XX OS Homo sapiens.
XX PN WO2003104277-A2.
XX PD 18-DEC-2003.
XX PF 05-JUN-2003; 2003WO-JP007123.
XX PR 05-JUN-2002; 2002JP-00164257.
XX PR 06-JUN-2002; 2002US-0385912P.
XX PR 26-DEC-2002; 2002JP-00377326.
XX PR 27-DEC-2002; 2002US-0436467P.
XX PR 15-MAY-2003; 2003JP-00137505.
XX PR 16-MAY-2003; 2003US-0470836P.
XX PA (ASAH ) ASAH KASEI KK.

```

XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
PI
XX WPI; 2004-122214/12.
DR P-PSDB; ADI26163.
XX
XX New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX
XX Claim 4; SEQ ID NO 127; 1368pp; English.
XX
XX The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficiently promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.
XX
SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 12.2 Length: 2361
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADI26162 (1-2361)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 666 AATCTGACCGAGGTGCCACGGACCTG 692

RESULT 26
ADL30833
ID ADL30833 standard; cDNA; 2361 BP.
XX
XX
AC ADL30833;
XX
XX 20-MAY-2004 (first entry)
XX
XX Full length human cDNA clone SeqID 2866.
DE
XX human; medicine; signal transduction; glycoprotein; transcription;
KW oligo-capping method; ss; gene.
XX
XX Homo sapiens.
OS
XX
XX EP1396543-A2.
PN
XX
XX 10-MAR-2004.
PD
XX
XX 07-JUL-2000; 2003EP-00025638.
PF
XX

PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2004-204755/20.
DR P-PSDB; ADL30834.
XX
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX
XX Example 1; SEQ ID NO 2866; 1340pp; English.
XX
XX This invention relates to a novel primers useful for synthesizing full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX
SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 12.2 Length: 2361
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADL30833 (1-2361)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 666 AATCTGACCGAGGTGCCACGGACCTG 692

RESULT 27
ABX40605
ID ABX40605 standard; cDNA; 343 BP.
XX
XX ABX40605;
XX
XX 20-FEB-2003 (first entry)
XX
XX Bovine EST associated with lactation/muscle/fat deposition #5770.
DE
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
XX
XX Bos Taurus.
OS
XX
XX US2002137139-A1.
PN
XX
XX 26-SEP-2002.
PD
XX
XX 24-SEP-2001; 2001US-00960352.
PF
XX
XX 12-JAN-1999; 99US-0115707P.
PR
XX 11-JAN-2000; 2000US-00480902.
PR
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
PA

XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
DR
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 5770; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 343 BP; 40 A; 146 C; 108 G; 49 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 16 Length: 343
Score: 41.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.1% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x ABX40605 (1-343)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 131 AACCTGACCGAGGTGCGCGGACCTG 157

RESULT 28
ABK87175
ID ABK87175 standard; cDNA; 1260 BP.
XX
XX ABK87175;
XX
DT 07-OCT-2002 (first entry)
XX
XX cDNA encoding feline oncofetal leucine-rich glycoprotein, 574.
DE
XX Feline; cat; oncofetal leucine-rich glycoprotein; 574; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX
XX Felis sp.
OS
XX Key Location/Qualifiers
FH 1. .1260
PT CDS

PT
XX
XX WO200238612-A2.
XX
XX 16-MAY-2002.
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Myers K, Drury N, Carroll M;
XX WPI; 2002-557449/59.
DR P-PSDB; AAU98694.
XX
XX Novel canine or feline 574 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 4; Page 68; 68pp; English.
XX
XX The present invention relates to the isolation of canine and feline
CC oncofetal leucine-rich glycoproteins known as 574, and the
CC polynucleotide sequences encoding them. The 574 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 574 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 574
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC calling foetal abnormalities, or to determine the sex of the foetus(es).
XX The present sequence encodes feline 574 protein

SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 70.2 Length: 1260
Score: 41.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.1% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x ABK87175 (1-1260)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 238 AACCTGACCGAGGTGCGCGGACCTG 264

RESULT 29
ADB97513
ID ADB97513 standard; DNA; 1260 BP.
XX
XX ADB97513;
XX
DT 04-DEC-2003 (first entry)
XX
XX Feline 574 antigen DNA.
XX
XX Major Histocompatibility Complex class I peptide epitope; MHC;
KW 574 antigen; 574 epitope; polypeptide string; vaccine; T cell;
KW cytostatic; cancer; feline; gene; ds.
XX
XX Unidentified.
OS

```

XX FH Key Location/Qualifiers
XX CDS 1..1260
XX FT /*tag= a
XX FT /product= "Feline 5T4 antigen protein"
XX PN WO2003068816-A1.
XX PD 21-AUG-2003.
XX PF 13-FEB-2003; 2003WO-GB000670.
XX PR 13-FEB-2002; 2002GB-00003419.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll M, Kingsman S, Redchenko I;
XX DR WPI; 2003-637141/60.
XX DR P-PSDB; ADB97520.
XX PT New major histocompatibility complex class I peptide epitopes from human
XX PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
XX PT disease, particularly cancer.
XX PS Disclosure; Page 67; 73pp; English.
XX CC The invention relates to a novel Major Histocompatibility Complex (MHC)
XX CC class I peptide epitope of the 5T4 antigen. The invention further
XX CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
XX CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
XX CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
XX CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
XX CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
XX CC comprising the above; a method for treating and/or preventing a disease
XX CC in a subject by administering the vaccine; an agent capable of binding
XX CC specifically to the 5T4 epitope and/or its encoding nucleic acid; a method
XX CC comprising detecting the presence of the 5T4 epitope or its encoding
XX CC nucleic acid in a subject; and a T cell line or clone capable of
XX CC specifically recognising the 5T4 epitope in conjunction with an MHC class
XX CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
XX CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
XX CC system or cell is useful in the prevention and/or treatment of a disease,
XX CC particularly cancer. The detection method is useful for diagnosing or
XX CC monitoring the progression of a cancerous disease, and for detecting the
XX CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
XX CC is useful in the manufacture of a medicament for treating and/or
XX CC preventing a disease. This polynucleotide sequence represents the feline
XX CC 5T4 antigen coding DNA of the invention.
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 70.2 Length: 1260
Score: 41.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.1% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-11 (1-9) x ADB97513 (1-1260)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 238 AACCTGACCGAGGTGCCCGGACCTG 264

RESULT 30
ADB97452
ID ADB97452 standard; DNA; 1260 BP.
XX AC ADB97452;
XX DT 04-DEC-2003 (first entry)

XX FH Key Location/Qualifiers
XX CDS 1..1260
XX FT /*tag= a
XX FT /product= "Feline 5T4 antigen protein"
XX PN WO2003068815-A2.
XX PD 21-AUG-2003.
XX PF 13-FEB-2003; 2003WO-GB000618.
XX PR 13-FEB-2002; 2002GB-00003420.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll M, Harrop R, Kingsman S;
XX DR WPI; 2003-663795/62.
XX DR P-PSDB; ADB97455.
XX PT New Major Histocompatibility Complex class II peptide epitope of 5T4,
XX PT useful for manufacturing a medicament for diagnosing, preventing and/or
XX PT treating a disease, e.g. cancer.
XX PS Disclosure; Page 49; 63pp; English.
XX CC The invention relates to a Major Histocompatibility Complex (MHC) class
XX CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
XX CC clone has a cytostatic activity, as it is useful in manufacturing a
XX CC medicament for preventing and/or treating a disease, particularly cancer.
XX CC The methods are useful for detecting T-cells capable of specifically
XX CC recognising a peptide epitope in conjunction with an MHC molecule, for
XX CC diagnosing or monitoring the progression of a cancerous disease, or for
XX CC detecting the presence of a peptide or nucleic acid using an agent. The
XX CC MHC class II peptide epitope of the invention can be used in gene therapy
XX CC or as part of a vaccine. This polynucleotide sequence represents the DNA
XX CC coding for the feline 5T4 protein.
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 70.2 Length: 1260
Score: 41.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.1% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-11 (1-9) x ADB97452 (1-1260)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 238 AACCTGACCGAGGTGCCCGGACCTG 264

RESULT 31
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX AC AAF89736;
XX DT 23-JUL-2001 (first entry)
XX DE Nucleotide sequence of canine 5T4 protein.
XX KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;

```

KW hypersensitivity; autoimmune disease; central nervous system disorder;
 KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
 KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
 KW Helicobacter-related disease; immune disorder; ss.
 XX
 OS Canis sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1263
 FT /*tag= a
 FT /product= "574"
 XX
 PN WO200136486-A2.
 XX
 PD 25-MAY-2001.
 XX
 XX 13-NOV-2000; 2000WO-GB004317.
 PF
 XX 18-NOV-1999; 99WO-GB003859.
 PR
 XX 15-FEB-2000; 2000GB-00003527.
 PR
 XX 02-MAR-2000; 2000GB-00005071.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Kingman A, Kingman SM, Bebbington CR, Carroll MW, Ellard FM;
 PI Myers KA;
 XX
 DR WPI, 2001-343805/36.
 DR P-PSDB; AAB83839.
 XX
 XX Use of single chain antibody capable of recognizing a disease associated
 PT molecule for manufacturing a medicament for preventing and/or treating a
 PT disease condition associated with disease associated molecule.
 XX
 PS Disclosure; Fig 26; 118pp; English.
 XX
 CC The specification describes the use of a single chain antibody (ScFv),
 CC which is capable of recognizing a disease associated molecule in the
 CC manufacture of a medicament for the prevention and treatment of a disease
 CC condition. The ScFv antibody is useful in the manufacture of a
 CC medicament, for affecting a disease in vivo, for preparing a
 CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
 CC treatment of a disease. The ScFv antibody is also useful for treating
 CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
 CC diseases, cancers, central nervous system disorders including Parkinson's
 CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
 CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
 CC related diseases, and other immune disorders. The present sequence
 CC encodes a 574 protein, which is used to produce ScFv of the invention
 XX
 SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 70.4 Length: 1263
 Score: 41.00 Matches: 8
 Percent Similarity: 88.9% Conservativity: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 89.1% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-11 (1-9) x AAF89736 (1-1263)
 QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 241 AACCTGACCGAGGTGCCGCGACCTG 267
 RESULT 32
 ABK87174
 ID ABK87174 standard; cDNA; 1263 BP.
 XX
 AC ABK87174;
 XX
 DT 07-OCT-2002 (first entry)

XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 574.
 DE
 XX Canine; dog; oncofoetal leucine-rich glycoprotein; 574; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX
 OS Canis sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1263
 FT /*tag= a
 FT /product= "574 protein"
 XX
 PN WO200238612-A2.
 XX
 PD 16-MAY-2002.
 XX
 XX 13-NOV-2001; 2001WO-GB005004.
 PF
 XX 13-NOV-2000; 2000WO-GB004317.
 PR
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 PA
 XX Myers K, Drury N, Carroll M;
 XX WPI, 2002-557449/59.
 DR P-PSDB; AAU98693.
 XX
 XX Novel canine or feline 574 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.
 XX
 PS Claim 1; Page 67; 68pp; English.
 XX
 CC The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 574, and the
 CC polynucleotide sequences encoding them. The 574 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 574 in extracts of plasma, urine, 574
 CC tissues, and in cell culture media. Antibodies specific for the 574
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes canine 574 protein
 XX
 SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 70.4 Length: 1263
 Score: 41.00 Matches: 8
 Percent Similarity: 88.9% Conservativity: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 89.1% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-11 (1-9) x ABK87174 (1-1263)
 QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 241 AACCTGACCGAGGTGCCGCGACCTG 267
 RESULT 33
 ADN46845 12/c
 Continuation (13 of 21) of ADN46845 from base 1200001 (Thermococcus kodakaraensis KOD1

WP Sequence split into 21 fragments LOCUS ADN46845 Accession Adn46845

```
WP Fragment Name      Begin      End
WP ADN46845_00        1        110000
WP ADN46845_01       100001    210000
WP ADN46845_02       200001    310000
WP ADN46845_03       300001    410000
WP ADN46845_04       400001    510000
WP ADN46845_05       500001    610000
WP ADN46845_06       600001    710000
WP ADN46845_07       700001    810000
WP ADN46845_08       800001    910000
WP ADN46845_09       900001   1010000
WP ADN46845_10      1000001   1110000
WP ADN46845_11      1100001   1210000
WP ADN46845_12      1200001   1310000
WP ADN46845_13      1300001   1410000
WP ADN46845_14      1400001   1510000
WP ADN46845_15      1500001   1610000
WP ADN46845_16      1600001   1710000
WP ADN46845_17      1700001   1810000
WP ADN46845_18      1800001   1910000
WP ADN46845_19      1900001   2010000
WP ADN46845_20      2000001   2089378
```

Alignment Scores:

Pred. No.:	1.13e+04	Length:	110000
Score:	41.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	89.1%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-11 (1-9) x ADN46845_12 (1-110000)

```
Oy 1 AsnLeuThrGluValProThrAspLeu 9
    |||::|||::|||::|||::|||::|||
Db 58561 AATATCACCAGGTGCCACTGACCTC 58535
```

RESULT 34

ADN47591_08
Continuation (9 of 21) of ADN47591 from base 800001 (Thermococcus kodakaraensis KOD1 gen
WP Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591

```
WP Fragment Name      Begin      End
WP ADN47591_00        1        110000
WP ADN47591_01       100001    210000
WP ADN47591_02       200001    310000
WP ADN47591_03       300001    410000
WP ADN47591_04       400001    510000
WP ADN47591_05       500001    610000
WP ADN47591_06       600001    710000
WP ADN47591_07       700001    810000
WP ADN47591_08       800001    910000
WP ADN47591_09       900001   1010000
WP ADN47591_10      1000001   1110000
WP ADN47591_11      1100001   1210000
WP ADN47591_12      1200001   1310000
WP ADN47591_13      1300001   1410000
WP ADN47591_14      1400001   1510000
WP ADN47591_15      1500001   1610000
WP ADN47591_16      1600001   1710000
WP ADN47591_17      1700001   1810000
WP ADN47591_18      1800001   1910000
WP ADN47591_19      1900001   2010000
WP ADN47591_20      2000001   2089378
```

Alignment Scores:

Pred. No.:	1.13e+04	Length:	110000
Score:	41.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	89.1%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-11 (1-9) x ADN47591_08 (1-110000)

```
Oy 1 AsnLeuThrGluValProThrAspLeu 9
    |||::|||::|||::|||::|||::|||
Db 30817 AATATCACCAGGTGCCACTGACCTC 30843
```

RESULT 35

ADN46123_12/c
Continuation (13 of 21) of ADN46123 from base 1200001 (Thermococcus kodakaraensis KOD1 gen
WP Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123

```
WP Fragment Name      Begin      End
WP ADN46123_00        1        110000
WP ADN46123_01       100001    210000
WP ADN46123_02       200001    310000
WP ADN46123_03       300001    410000
WP ADN46123_04       400001    510000
WP ADN46123_05       500001    610000
WP ADN46123_06       600001    710000
WP ADN46123_07       700001    810000
WP ADN46123_08       800001    910000
WP ADN46123_09       900001   1010000
WP ADN46123_10      1000001   1110000
WP ADN46123_11      1100001   1210000
WP ADN46123_12      1200001   1310000
WP ADN46123_13      1300001   1410000
WP ADN46123_14      1400001   1510000
WP ADN46123_15      1500001   1610000
WP ADN46123_16      1600001   1710000
WP ADN46123_17      1700001   1810000
WP ADN46123_18      1800001   1910000
WP ADN46123_19      1900001   2010000
WP ADN46123_20      2000001   2089378
```

Alignment Scores:

Pred. No.:	1.13e+04	Length:	110000
Score:	41.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	89.1%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-11 (1-9) x ADN46123_12 (1-110000)

```
Oy 1 AsnLeuThrGluValProThrAspLeu 9
    |||::|||::|||::|||::|||::|||
Db 58561 AATATCACCAGGTGCCACTGACCTC 58535
```

RESULT 36

ADN47209_08
Continuation (9 of 21) of ADN47209 from base 800001 (Thermococcus kodakaraensis KOD1 gen
WP Sequence split into 21 fragments LOCUS ADN47209 Accession Adn47209

```
WP Fragment Name      Begin      End
WP ADN47209_00        1        110000
WP ADN47209_01       100001    210000
WP ADN47209_02       200001    310000
WP ADN47209_03       300001    410000
WP ADN47209_04       400001    510000
WP ADN47209_05       500001    610000
WP ADN47209_06       600001    710000
WP ADN47209_07       700001    810000
WP ADN47209_08       800001    910000
WP ADN47209_09       900001   1010000
WP ADN47209_10      1000001   1110000
WP ADN47209_11      1100001   1210000
WP ADN47209_12      1200001   1310000
WP ADN47209_13      1300001   1410000
WP ADN47209_14      1400001   1510000
WP ADN47209_15      1500001   1610000
WP ADN47209_16      1600001   1710000
WP ADN47209_17      1700001   1810000
WP ADN47209_18      1800001   1910000
WP ADN47209_19      1900001   2010000
WP ADN47209_20      2000001   2089378
```


Alignment Scores:
Pred. No.: 1.13e+04 Length: 110000
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.1% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADN47209_08 (1-110000)

QY 1 AsnLeuThrGluValProThrAspLeu 9

DB 30817 AATATCACCAGGTGCCCACTGACCTC 30843

RESULT 37

ADN46464_12/c Continuation (13 of 21) of ADN46464 from base 1200001 (Thermococcus kodakaraensis KOD1)

WP	Sequence split into 21 fragments	LOCUS	ADN46464	Accession	ADN46464
WP	Fragment Name	Begin	End		
WP	ADN46464_00	1	110000		
WP	ADN46464_01	100001	210000		
WP	ADN46464_02	200001	310000		
WP	ADN46464_03	300001	410000		
WP	ADN46464_04	400001	510000		
WP	ADN46464_05	500001	610000		
WP	ADN46464_06	600001	710000		
WP	ADN46464_07	700001	810000		
WP	ADN46464_08	800001	910000		
WP	ADN46464_09	900001	1010000		
WP	ADN46464_10	1000001	1110000		
WP	ADN46464_11	1100001	1210000		
WP	ADN46464_12	1200001	1310000		
WP	ADN46464_13	1300001	1410000		
WP	ADN46464_14	1400001	1510000		
WP	ADN46464_15	1500001	1610000		
WP	ADN46464_16	1600001	1710000		
WP	ADN46464_17	1700001	1810000		
WP	ADN46464_18	1800001	1910000		
WP	ADN46464_19	1900001	2010000		
WP	ADN46464_20	2000001	2089378		

Alignment Scores:
Pred. No.: 1.13e+04 Length: 110000
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.1% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADN46464_12 (1-110000)

QY 1 AsnLeuThrGluValProThrAspLeu 9

DB 58561 AATATCACCAGGTGCCCACTGACCTC 58535

RESULT 38

ADN47960_08 Continuation (9 of 21) of ADN47960 from base 800001 (Thermococcus kodakaraensis KOD1)

WP	Sequence split into 21 fragments	LOCUS	ADN47960	Accession	ADN47960
WP	Fragment Name	Begin	End		
WP	ADN47960_00	1	110000		
WP	ADN47960_01	100001	210000		
WP	ADN47960_02	200001	310000		
WP	ADN47960_03	300001	410000		
WP	ADN47960_04	400001	510000		
WP	ADN47960_05	500001	610000		
WP	ADN47960_06	600001	710000		
WP	ADN47960_07	700001	810000		
WP	ADN47960_08	800001	910000		
WP	ADN47960_09	900001	1010000		
WP	ADN47960_10	1000001	1110000		
WP	ADN47960_11	1100001	1210000		

WP ADN47960_12 1200001 1310000
WP ADN47960_13 1300001 1410000
WP ADN47960_14 1400001 1510000
WP ADN47960_15 1500001 1610000
WP ADN47960_16 1600001 1710000
WP ADN47960_17 1700001 1810000
WP ADN47960_18 1800001 1910000
WP ADN47960_19 1900001 2010000
WP ADN47960_20 2000001 2089378

Alignment Scores:
Pred. No.: 1.13e+04 Length: 110000
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.1% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADN47960_08 (1-110000)

QY 1 AsnLeuThrGluValProThrAspLeu 9

DB 30817 AATATCACCAGGTGCCCACTGACCTC 30843

RESULT 39

ABQ42606/c
ID ABQ42606 standard; DNA; 792 BP.
XX ABQ42606;
AC ABQ42606;
XX 12-JUL-2002 (first entry)
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 29197.
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
OS Homo sapiens.
XX MO200218632-A2.
XX 07-MAR-2002.
XX 01-SEP-2001; 2001WO-EP010074.
XX 01-SEP-2000; 2000DE-01043826.
XX 05-SEP-2000; 2000DE-01044543.
XX (BPIG-) EPIGENOMICS AG.

Olek A, Piepenbrock C, Berlin K, Guetig D;

WPI; 2002-371829/40.

Determining the degree of cytosine methylation in genomic DNA, useful for diagnosis and prognosis, comprises selective hybridization of amplicons from chemically treated DNA.

Claim 12; 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs

CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and, respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

XX SQ Sequence 792 BP; 119 A; 94 C; 297 G; 282 T; 0 U; 0 Other;

Alignment Scores: 67.8 Length: 792
 Pred. No.: 40.00 Matches: 7
 Score: 100.0% Conservatives: 2
 Percent Similarity: 77.8% Mismatches: 0
 Best Local Similarity: 87.0% Indels: 0
 Query Match: 6 Gaps: 0

US-10-774-176-11 (1-9) x ABQ42606 (1-792)

QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 208 AATCTAACCGAATATACCCAGACCTA 182

RESULT 40

ABQ42607

ID ABQ42607 standard; DNA; 792 BP.

AC ABQ42607;

DT 12-JUL-2002 (first entry)

DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29198.

KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

OS Homo sapiens.

PN WO200218632-A2.

XX 07-MAR-2002.

PF 01-SEP-2001; 2001WO-EP010074.

XX 01-SEP-2000; 2000DE-01043826.

PR 05-SEP-2000; 2000DE-01044543.

XX (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridisation to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridised to the two classes of

CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

XX SQ Sequence 792 BP; 282 A; 297 C; 94 G; 119 T; 0 U; 0 Other;

Alignment Scores: 67.8 Length: 792
 Pred. No.: 40.00 Matches: 7
 Score: 100.0% Conservatives: 2
 Percent Similarity: 77.8% Mismatches: 0
 Best Local Similarity: 87.0% Indels: 0
 Query Match: 6 Gaps: 0

US-10-774-176-11 (1-9) x ABQ42607 (1-792)

QY 1 AsnLeuThrGluValProThrAspLeu 9

DB 585 AATCTAACCGAATATACCCAGACCTA 611

RESULT 41

ABQ42584/C

ID ABQ42584 standard; DNA; 795 BP.

XX AC ABQ42584;

XX 12-JUL-2002 (first entry)

DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 29175.

KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

OS Homo sapiens.

PN WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP010074.

XX 01-SEP-2000; 2000DE-01043826.

PR 05-SEP-2000; 2000DE-01044543.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the

Alignment Scores:	
Pred. No.:	95
Length:	691

Score: 39.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 84.8% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADG45384 (1-691)

QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 430 AATCTTAGAGAGATTCCTACAGATCTT 404

RESULT 46

ID ADH22691/C
 ADH22691 standard; DNA; 691 BP.

XX AC ADH22691;
 XX DT 11-MAR-2004 (first entry)
 XX DE Partial DNA sequence of a rat kidney toxicity predictive gene (19).

XX KW kidney toxicity; toxicology; predictive model; gene expression profile;
 XX KW toxic damage; kidney tubule necrosis; acute renal failure; rat; db.
 XX OS Rattus sp.
 XX PN WO2003100030-A2.
 XX PD 04-DEC-2003.
 XX PF 27-FEB-2003; 2003WO-US006196.
 XX PR 27-FEB-2002; 2002US-0361128P.
 XX PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
 XX PI Kier L, Nolan TD, Sankar U, Derbel M;
 XX WPI; 2004-035137/03.

XX PT Predicting kidney toxicity in an individual to an agent, useful for
 PT predicting toxic responses to one or more agents comprising measuring the
 PT expression of one or more kidney toxicity predictive genes.
 XX Claim 1; Page 319; 389pp; English.
 XX This invention relates to novel isolated kidney toxicity predictive genes
 CC and methods of using such genes in the field of toxicology. Specifically,
 CC it refers to novel genes that can be used for generating predictive
 CC models, which in turn are useful for predicting the in vivo toxic
 CC response to one or more agents. The present invention describes obtaining
 CC a gene expression profile from a biological sample and using the
 CC predictive model to determine whether an agent will induce kidney
 CC toxicity in the individual. As such, it can be used to detect any toxic
 CC effects that may be manifested as long lasting or chronic consequences
 CC including irreversible toxicity or carcinogenesis. Furthermore, the
 CC predictive genes can be considered as therapeutic targets for toxic
 CC damage or to ameliorate specific disease conditions such as kidney tubule
 CC necrosis or acute renal failure, as well as for additional screening
 CC assays. This polynucleotide sequence is the partial gene sequence of a
 CC rat kidney toxicity predictive gene of the invention.

XX SQ Sequence 691 BP; 168 A; 177 C; 157 G; 185 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 95 Length: 691
 Score: 39.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 84.8% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-11 (1-9) x ADH22691 (1-691)

QY 1 AsnLeuThrGluValProThrAspLeu 9
 DB 430 AATCTTAGAGAGATTCCTACAGATCTT 404

RESULT 47

ID ADH91053/C
 ADH91053 standard; CDNA; 691 BP.

XX AC ADH91053;
 XX DT 16-DEC-2004 (first entry)
 XX DE Spleen necrosis predictive cDNA sequence, SEQ ID No 9.

XX KW altered expression; toxic response; spleen; toxicity; lymphoid; gene; ss.
 XX OS Unidentified.
 XX PN WO2004083402-A2.
 XX PD 30-SEP-2004.
 XX PF 17-MAR-2004; 2004WO-US008371.
 XX PR 17-MAR-2003; 2003US-0455443P.
 XX PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
 XX PI Sankar U, Kier L, Derbel M, Nolan T;
 XX WPI; 2004-691048/67.
 XX PT New composition comprises cDNAs useful for detecting altered expression
 PT of genes in a toxic response of the spleen or for predicting toxic
 PT responses to one or more agents including lymphoid tissue types or other
 PT species.

XX Claim 1; SEQ ID NO 9; 249pp; English.

XX The invention relates to a novel composition comprising cDNAs for use in
 CC detecting the altered expression of genes in a toxic response of the
 CC spleen, where the cDNAs comprises 50-816 base pairs (ADR91045-ADR91348)
 CC or their complete complements. The invention further comprises:
 CC monitoring the treatment of compound toxicity in a sample; predicting the
 CC spleen toxicity in an individual to an agent; predicting the spleen
 CC toxicity of an agent using an in vitro system; a computer program product
 CC for predicting spleen toxicity from an expression profile of nucleic
 CC acids from a sample under test, comprising a computer readable medium
 CC bearing an encrypted training data set, encrypted lists of genes selected
 CC from the cDNAs, and a predictive model for causing a general purpose
 CC computer to predict the spleen toxicity of the sample based upon the
 CC training data set, the list of genes selected from the cDNAs, and the
 CC expression profile of nucleic acids from the sample; and an integrated
 CC system for predicting spleen toxicity, comprising means for measuring
 CC gene expression profiles of spleen predictive genes from samples exposed
 CC to the test agent and a computer system operably linked to the means that
 CC is capable of implementing a predictive model. The composition comprising
 CC cDNAs is useful for detecting altered expression of genes in a toxic
 CC response of the spleen or for predicting toxic responses to one or more
 CC agents including lymphoid tissue types or other species. The predictive
 CC genes and models of the invention are useful for identifying and
 CC evaluating various in vitro systems that can be used to accurately
 CC predict in vivo toxicity. This polynucleotide sequence represents one of
 CC the genes predictive for spleen necrosis of the invention.

XX SQ Sequence 691 BP; 168 A; 177 C; 157 G; 185 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 95 Length: 691
 Score: 39.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1

```
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-11 (1-9) x ADR91053 (1-691)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 430 AATCTTAGAGAGATTCCTACAGATCTT 404

RESULT 48
ABK79398
ID ABK79398 standard; DNA; 933 BP.
XX
AC ABK79398;
XX
DT 13-AUG-2002 (first entry)
DE Bacillus clausii genomic sequence tag (GST) #2241.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus clausii.
XX
FN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
PR 06-OCT-2000; 2000US-00680598.
PR 27-MAR-2001; 2001US-0279526P.
XX
PA (NOVO ) NOVOZYMES BIOTECH INC.
PA (NOVO ) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
PS Claim 11; SEQ ID NO 6689; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 933 BP; 267 A; 219 C; 235 G; 212 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 134 Length: 933
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x ABK79398 (1-933)
QY 1 AsnLeuThrGluValProThrAspLeu 9
DB 669 AACTTGACTGAATGTCCCACTGACATC 695

RESULT 49
ABK63515
ID ABK63515 standard; cDNA; 2977 BP.
XX
AC ABK63515;
XX
DT 18-JUN-2002 (first entry)
DE Rat sequence differentially expressed in response to a hepatotoxin #1422.
XX
KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW differential expression; centrilobular necrosis; steatosis.
XX
OS Rattus norvegicus.
XX
FN WO200210453-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US023872.
XX
PR 31-JUL-2000; 2000US-0222040P.
PR 02-NOV-2000; 2000US-0244880P.
PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Blashoff MR;
XX
DR WPI; 2002-241625/29.
XX
PT Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or cells
PT exposed to the toxin and comparing these to gene expression in unexposed
PT tissues or cells.
XX
PS Claim 1; SEQ ID NO 1422; 239pp; English.
XX
CC The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic effect
CC of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression. The
CC method can also be used to identify an agent which modulates the toxic
CC response and predict cellular pathways that a compound modulates in a
CC cell. The methods utilise a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridises to a gene listed in the specification, a computer
```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame plus p2n model.

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignment)
171.290 Million cell updates/sec

Title: US-10-774-176-11

Perfect score: 46

Sequence: 1 NLTEVPTDL 9

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abse/ABSSWEB.spool/US10774176/runat_24042006_165114_19197/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEADING=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs804
-USER=US10774176 @CGN 1.1 6765 @runat_24042006_165114_19197 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sy.*
12: gb.un.*
13: gb.vi.*
14: gb.htg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	46	100.0	927	6	AX829164	Sequence
2	46	100.0	1263	6	BD249731	Polypepti
3	46	100.0	1263	6	AX025011	Sequence

46	100.0	1263	6	AX316086	AX316086 Sequence
46	100.0	2053	6	CQ731678	CQ731678 Sequence
46	100.0	2053	8	H5ST40A	Z29083 Homo sapien
46	100.0	2359	6	BD127282	BD127282 Primer fo
46	100.0	2359	6	CQ782724	CQ782724 Sequence
46	100.0	2359	8	AK074786	AK074786 Homo sapi
46	100.0	2361	6	BD127283	BD127283 Primer fo
46	100.0	2361	6	CQ782726	CQ782726 Sequence
46	100.0	2361	6	AX961916	AX961916 Sequence
46	100.0	2361	8	AK074790	AK074790 Homo sapi
46	100.0	2379	8	BC037161	BC037161 Homo sapi
46	100.0	2714	8	AB168308	AB168308 Macaca fa
46	100.0	2551	8	HSJ492P14	HSJ492P14 Homo sapi
46	100.0	121909	8	HSJ492P14	AL121977 Human DNA
41	89.1	1260	6	AX467373	AX467373 Sequence
41	89.1	1260	6	AX821533	AX821533 Sequence
41	89.1	1260	6	AX821548	AX821548 Sequence
41	89.1	1263	6	AX149553	AX149553 Sequence
41	89.1	1263	6	AX467371	AX467371 Sequence
41	89.1	110000	1	AP006878.12	Continuation (13 o
41	89.1	187783	14	AC137462	AC137462 Rattus no
41	89.1	225405	14	AC158075	AC158075 Bos tauru
41	89.1	241280	14	AC115158	AC115158 Rattus no
41	89.1	267375	14	AC095339	AC095339 Rattus no
40	87.0	701	10	BV561288	BV561288 ggi26e11.
40	87.0	804	10	BV564129	BV564129 gq668d12.
40	87.0	35413	2	AF039719	AF039719 Caenorhab
40	87.0	106590	14	AC151952	AC151952 Pan trogl
40	87.0	166980	14	AC166203	AC166203 Oryctolag
40	87.0	180623	8	AC092341	AC092341 Homo sapi
40	87.0	181161	14	AC092347	AC092347 Homo sapi
40	87.0	188254	14	AC166207	AC166207 Oryctolag
40	87.0	191684	8	AC146469	AC146469 Pan trogl
40	87.0	231264	14	CR388209	CR388209 Danio rer
38	84.8	394	10	BV090626	BV090626 RPAMMSEQ
39	84.8	691	6	AX525488	AX525488 Sequence
39	84.8	933	6	AX438274	AX438274 Sequence
39	84.8	2949	5	AJ720394	AJ720394 Gallus ga
39	84.8	2977	9	AX401746	AX401746 Sequence
39	84.8	3007	9	RATMAD	M73714 Rat micros
39	84.8	30916	15	BC003797	BC003797 Mus muscu
39	84.8	32478	9	AL672172	AL672172 Mouse cor
39	84.8	87118	8	AL355372	AL355372 Human DNA
39	84.8	108661	8	AC026736	AC026736 Homo sapi
39	84.8	110000	1	AP006627.40	Continuation (41 o
39	84.8	110000	15	AB017353.2	Continuation (3 of
39	84.8	111081	14	AC010249	AC010249 Homo sapi
39	84.8	133375	14	CR936285	CR936285 Danio rer
39	84.8	135439	5	BX571773	BX571773 Zebrafish
39	84.8	137165	14	AC011264	AC011264 Homo sapi
39	84.8	147599	9	AC159738	AC159738 Mus muscu
39	84.8	148833	14	AC155960	AC155960 Xenopus t
39	84.8	158802	5	CR9333001	CR9333001 Zebrafish
39	84.8	162918	8	AC026799	AC026799 Homo sapi
39	84.8	181566	9	AC127247	AC127247 Mus muscu
39	84.8	192202	14	AL591106	AL591106 Homo sapi
39	84.8	199784	14	AC118419	AC118419 Rattus no
39	84.8	202046	14	AC151872	AC151872 Lemur cat
39	84.8	204686	14	CR933526	CR933526 Danio rer
39	84.8	206294	9	AC108394	AC108394 Mus muscu
39	84.8	218558	14	AC025964	AC025964 Mus muscu
39	84.8	227722	14	AC107139	AC107139 Rattus no
39	84.8	240536	9	AC025910	AC025910 Mus muscu
39	84.8	244814	14	AC098223	AC098223 Rattus no
39	84.8	248573	14	AC096467	AC096467 Rattus no
39	84.8	255474	14	AC094934	AC094934 Rattus no
39	84.8	275197	14	AC095358	AC095358 Rattus no
38	82.6	90077	15	AP004915	AP004915 Lotus cor
38	82.6	93772	5	AL603715	AL603715 Zebrafish
38	82.6	100922	14	AP008179	AP008179 Lotus cor
38	82.6	103750	8	AL390246	AL390246 Human DNA
38	82.6	115247	14	AP007423	AP007423 Lotus cor

77	38	82.6	142080	14	AC069260	AC069260 Homo sapi	C 150	37	80.4	115119	15	AC155345	AC155345 Brassica
78	38	82.6	144356	8	AC021486	AC021486 Homo sapi	151	37	80.4	123661	5	BR927378	BR927378 Zebraphish
79	38	82.6	148643	14	AC011871	AC011871 Homo sapi	152	37	80.4	126368	14	CR774177	CR774177 Danio rer
80	38	82.6	151182	8	AL611942	AL611942 Human DNA	153	37	80.4	131309	14	AC156325	AC156325 Bos taurus
81	38	82.6	151552	14	AC023437	AC023437 Homo sapi	C 154	37	80.4	142314	8	AC098867	AC098867 Homo sapi
82	38	82.6	154233	14	AC108701	AC108701 Homo sapi	C 155	37	80.4	143039	8	AC016559	AC016559 Homo sapi
83	38	82.6	157929	9	AC147186	AC147186 Mus muscu	156	37	80.4	146026	5	BR248113	BR248113 Zebraphish
84	38	82.6	161508	14	AC092088	AC092088 Canis fam	157	37	80.4	152416	5	BR537336	BR537336 Zebraphish
85	38	82.6	163695	14	BR664733	BR664733 Homo sapi	C 158	37	80.4	153265	14	BR005346	BR005346 Danio rer
86	38	82.6	165203	8	AC093192	AC093192 Homo sapi	C 159	37	80.4	154456	5	CR450812	CR450812 Zebraphish
87	38	82.6	168659	14	AL359883	AL359883 Homo sapi	160	37	80.4	157047	14	CR855310	CR855310 Danio rer
88	38	82.6	169374	14	AC026050	AC026050 Homo sapi	C 161	37	80.4	158326	14	AC163338	AC163338 Mus muscu
89	38	82.6	172607	5	CR628410	CR628410 Zebraphish	C 162	37	80.4	161310	5	CR381708	CR381708 Zebraphish
90	38	82.6	172609	14	AP001463	AP001463 Homo sapi	C 163	37	80.4	163999	14	AC141207	AC141207 Rattus no
91	38	82.6	177003	14	AC022043	AC022043 Homo sapi	C 164	37	80.4	164436	14	AC051656	AC051656 Homo sapi
92	38	82.6	181070	14	AC145850	AC145850 Gallus ga	165	37	80.4	166257	14	AC165247	AC165247 Mus muscu
93	38	82.6	183203	14	AC078990	AC078990 Homo sapi	166	37	80.4	166681	5	BR322231	BR322231 Zebraphish
94	38	82.6	190253	14	AC067779	AC067779 Homo sapi	C 167	37	80.4	167381	5	BR571817	BR571817 Zebraphish
95	38	82.6	190723	8	AC069262	AC069262 Homo sapi	C 168	37	80.4	167492	14	AC164426	AC164426 Mus muscu
96	38	82.6	199715	5	CR376801	CR376801 Zebraphish	169	37	80.4	169912	9	AC121996	AC121996 Mus muscu
97	38	82.6	200300	14	CR547126	CR547126 Danio rer	C 170	37	80.4	170925	9	AC156398	AC156398 Mus muscu
98	38	82.6	201821	14	AC129105	AC129105 Homo sapi	C 171	37	80.4	172090	5	BR629340	BR629340 Zebraphish
99	38	82.6	202237	9	AC114825	AC114825 Mus muscu	C 172	37	80.4	172103	9	AC134575	AC134575 Mus muscu
100	38	82.6	203281	9	AC108435	AC108435 Mus muscu	C 173	37	80.4	172122	5	BR927315	BR927315 Zebraphish
101	38	82.6	205619	14	AC108275	AC108275 Rattus no	C 174	37	80.4	176239	5	BR248246	BR248246 Zebraphish
102	38	82.6	206168	9	AL672307	AL672307 Mouse DNA	C 175	37	80.4	176916	5	AL845493	AL845493 Zebraphish
103	38	82.6	212118	14	AC157062	AC157062 Bos taurus	C 176	37	80.4	177325	8	AC138744	AC138744 Homo sapi
104	38	82.6	222854	9	AC136717	AC136717 Mus muscu	C 177	37	80.4	178295	14	AC021798	AC021798 Homo sapi
105	38	82.6	233878	14	AC098146	AC098146 Rattus no	C 178	37	80.4	179137	14	CR847803	CR847803 Danio rer
106	38	82.6	237913	14	AC094676	AC094676 Rattus no	C 179	37	80.4	181043	14	AC138763	AC138763 Homo sapi
107	38	82.6	246109	14	AC111307	AC111307 Rattus no	C 180	37	80.4	181253	8	AC099065	AC099065 Mus muscu
108	38	82.6	248537	14	AC107098	AC107098 Rattus no	C 181	37	80.4	181373	9	AC150893	AC150893 Mus muscu
109	38	82.6	252365	14	AC123642	AC123642 Mus muscu	C 182	37	80.4	181547	9	AL512583	AL512583 Mouse DNA
110	38	82.6	253346	14	AC128440	AC128440 Rattus no	C 183	37	80.4	182147	8	AC062032	AC062032 Homo sapi
111	38	82.6	259722	14	AC134135	AC134135 Rattus no	C 184	37	80.4	183342	8	AC087490	AC087490 Homo sapi
112	38	82.6	339569	14	AC107306	AC107306 Homo sapi	C 185	37	80.4	183526	9	AC158642	AC158642 Mus muscu
113	37	80.4	611	5	OSEARP	M96154 Osmorus mor	186	37	80.4	184346	14	AC069138	AC069138 Homo sapi
114	37	80.4	615	10	BV077507	BV077507 S209P6423	187	37	80.4	184868	5	BR004967	BR004967 Zebraphish
115	37	80.4	640	10	BV337153	BV337153 S230P6326	188	37	80.4	186051	14	BR005324	BR005324 Mus muscu
116	37	80.4	650	10	BV214415	BV214415 S233P6384	189	37	80.4	188552	8	AC009994	AC009994 Homo sapi
117	37	80.4	652	4	AF075641	AF075641 Equus cab	C 190	37	80.4	189033	14	AC166006	AC166006 Colobus g
118	37	80.4	1169	2	AF470577	AF470577 Bolbophor	C 191	37	80.4	189411	8	AC096542	AC096542 Homo sapi
119	37	80.4	1169	2	AF470586	AF470586 Bolbophor	C 192	37	80.4	189962	14	AC148165	AC148165 Zee may
120	37	80.4	1169	2	AF470593	AF470593 Bolbophor	C 193	37	80.4	190817	9	AC110536	AC110536 Mus muscu
121	37	80.4	1169	2	AF470601	AF470601 Bolbophor	C 194	37	80.4	190998	5	BR649384	BR649384 Zebraphish
122	37	80.4	1232	2	AF470539	AF470539 Bolbophor	C 195	37	80.4	191699	14	AC108878	AC108878 Mus muscu
123	37	80.4	1233	2	AF470538	AF470538 Bolbophor	196	37	80.4	191805	8	AC027670	AC027670 Homo sapi
124	37	80.4	1233	2	AF470543	AF470543 Bolbophor	C 197	37	80.4	192929	14	CT025562	CT025562 Mus muscu
125	37	80.4	1233	2	AF470570	AF470570 Bolbophor	C 198	37	80.4	193064	9	AC124012	AC124012 Mus muscu
126	37	80.4	1233	2	AF470585	AF470585 Bolbophor	C 199	37	80.4	196155	5	BR247882	BR247882 Zebraphish
127	37	80.4	1233	2	AF470589	AF470589 Bolbophor	C 200	37	80.4	197401	9	AC153612	AC153612 Mus muscu
128	37	80.4	1233	2	AF470605	AF470605 Bolbophor	C 201	37	80.4	198825	5	BR293991	BR293991 Zebraphish
129	37	80.4	1233	2	AF470613	AF470613 Bolbophor	C 202	37	80.4	199109	9	AC163742	AC163742 Mus muscu
130	37	80.4	1234	2	AF470574	AF470574 Bolbophor	C 203	37	80.4	199853	9	AC164624	AC164624 Mus muscu
131	37	80.4	1234	2	AF470581	AF470581 Bolbophor	C 204	37	80.4	200574	9	BR005304	BR005304 Mouse DNA
132	37	80.4	1288	2	AY222172	AY222172 Ichthyoco	C 205	37	80.4	200967	9	AC083910	AC083910 Mus muscu
133	37	80.4	1360	2	AY222171	AY222171 Cardiocep	C 206	37	80.4	205428	14	CR936463	CR936463 Danio rer
134	37	80.4	1752	2	AY620907	AY620907 Horreolan	C 207	37	80.4	206030	14	AC021439	AC021439 Homo sapi
135	37	80.4	1881	6	AX413848	AX413848 Sequence	C 208	37	80.4	206606	14	AC158012	AC158012 Bos taurus
136	37	80.4	1881	6	AX415833	AX415833 Sequence	C 209	37	80.4	206709	14	CR762441	CR762441 Danio rer
137	37	80.4	1968	5	BC096939	BC096939 Danio rer	C 210	37	80.4	207008	14	CR847513	CR847513 Danio rer
138	37	80.4	2001	2	AKL12791	AKL12791 Ciona int	C 211	37	80.4	212931	14	AC157421	AC157421 Bos taurus
139	37	80.4	4618	8	AC079119	AC079119 Homo sapi	C 212	37	80.4	213217	14	AC140252	AC140252 Mus muscu
140	37	80.4	33455	2	CEW04A8	Z82069 Caenorhabdi	C 213	37	80.4	215027	14	CR759958	CR759958 Danio rer
141	37	80.4	42052	2	U80029	U80029 Caenorhabdi	C 214	37	80.4	216647	5	BR649529	BR649529 Zebraphish
142	37	80.4	64040	14	AC080026	AC080026 Homo sapi	C 215	37	80.4	221396	9	AC105487	AC105487 Rattus no
143	37	80.4	77634	5	CR391915	CR391915 Zebraphish	C 216	37	80.4	221478	14	AC079545	AC079545 Mus muscu
144	37	80.4	82356	14	AC144912	AC144912 Mus muscu	C 217	37	80.4	223578	5	AL953901	AL953901 Zebraphish
145	37	80.4	96747	14	CT009598	CT009598 Danio rer	C 218	37	80.4	223613	14	AC126216	AC126216 Rattus no
146	37	80.4	102151	14	AP007814	AP007814 Lotus cor	C 219	37	80.4	224985	14	AC130621	AC130621 Rattus no
147	37	80.4	109740	5	BR248322	BR248322 Zebraphish	C 220	37	80.4	225606	14	AC122078	AC122078 Rattus no
148	37	80.4	110000	1	CP000025	Continuation (3 of	221	37	80.4	228327	5	AL844514	AL844514 Zebraphish
149	37	80.4	110000	14	AC020850	Continuation (3 of	222	37	80.4	228767	5	BR005261	BR005261 Zebraphish

369	36	78.3 151300	5	CR925718	CR925718 Zebrafish	C 442	36	78.3 191176	9	AC121817	AC121817 Mus muscu
370	36	78.3 152725	15	AC120889	Oryza sat	443	36	78.3 191193	5	BX511113	BX511113 Zebrafish
371	36	78.3 152831	9	AC155649	Mus muscu	444	36	78.3 191823	8	AC146174	AC146174 Pan trogl
372	36	78.3 152888	9	AC121892	Mus muscu	C 445	36	78.3 191840	14	AC092656	AC092656 Homo sapi
373	36	78.3 155000	14	AC023689	Drosophil	446	36	78.3 191918	14	AC026484	AC026484 Homo sapi
374	36	78.3 156304	14	AC152919	Medicago	447	36	78.3 191928	14	CR385041	CR385041 Danio rer
375	36	78.3 156806	14	AC021153	AC021153 Homo sapi	448	36	78.3 192031	9	AL732504	AL732504 Mouse DNA
376	36	78.3 158073	8	AC069298	AC069298 Homo sapi	C 449	36	78.3 192304	5	AL954812	AL954812 Zebrafish
377	36	78.3 158223	14	BX977722	Danio rer	450	36	78.3 193168	5	BX000443	BX000443 Zebrafish
378	36	78.3 159077	14	AC027610	Homo sapi	451	36	78.3 194021	5	EX470266	EX470266 Zebrafish
379	36	78.3 159451	14	CR848665	Danio rer	452	36	78.3 194215	14	AC024422	AC024422 Homo sapi
380	36	78.3 159468	14	AC015280	AC015280 Drosophil	453	36	78.3 194866	14	AC142140	AC142140 Rattus no
381	36	78.3 160106	14	AC159924	AC159924 Orycolag	C 454	36	78.3 195366	14	CR847978	CR847978 Danio rer
382	36	78.3 160603	15	AC084818	AC084818 Oryza sat	455	36	78.3 196480	14	AC112230	AC112230 Homo sapi
383	36	78.3 161275	5	AL954129	AL954129 Zebrafish	456	36	78.3 197359	14	AC117549	AC117549 Mus muscu
384	36	78.3 162345	2	AC006415	AC006415 Drosophil	C 457	36	78.3 197412	9	AC074225	AC074225 Mus muscu
385	36	78.3 162640	2	AC154047	AC154047 Drosophil	458	36	78.3 197460	14	AC160817	AC160817 Bos tauru
386	36	78.3 163183	14	AC135306	AC135306 Rattus no	C 459	36	78.3 198853	14	AL365354	AL365354 Homo sapi
387	36	78.3 163300	5	CR388011	CR388011 Zebrafish	460	36	78.3 199831	5	BX470099	BX470099 Zebrafish
388	36	78.3 163413	14	AC118693	AC118693 Mus muscu	C 461	36	78.3 200253	9	AC162382	AC162382 Mus muscu
389	36	78.3 163717	9	AC116527	AC116527 Mus muscu	C 462	36	78.3 201526	14	CR388099	CR388099 Danio rer
390	36	78.3 163866	14	CR388127	CR388127 Danio rer	463	36	78.3 201542	9	AL7332483	AL7332483 Mouse DNA
391	36	78.3 164469	14	BX914209	BX914209 Danio rer	464	36	78.3 202872	5	BX248581	BX248581 Zebrafish
392	36	78.3 165019	8	AC138696	AC138696 Homo sapi	C 465	36	78.3 203705	5	CR293501	CR293501 Zebrafish
393	36	78.3 165531	4	CR853300	CR853300 Wallaby D	C 466	36	78.3 204190	14	AC140696	AC140696 Rattus no
394	36	78.3 165933	5	AL772362	AL772362 Zebrafish	C 467	36	78.3 204592	14	AC153458	AC153458 Bos tauru
395	36	78.3 166433	14	AC156890	AC156890 Bos tauru	C 468	36	78.3 204856	9	AC163642	AC163642 Mus muscu
396	36	78.3 167366	9	AC154496	AC154496 Mus muscu	C 469	36	78.3 206431	14	AL935191	AL935191 Danio rer
397	36	78.3 167396	5	CR387977	CR387977 Zebrafish	C 470	36	78.3 206438	5	AL935268	AL935268 Zebrafish
398	36	78.3 167806	14	AC092572	AC092572 Homo sapi	471	36	78.3 206592	9	AC122471	AC122471 Mus muscu
399	36	78.3 167927	14	AC162782	AC162782 Danio rer	C 472	36	78.3 208207	5	AL145450	AL145450 Mus muscu
400	36	78.3 168375	14	AC158881	AC158881 Bos tauru	473	36	78.3 208746	5	CR387978	CR387978 Zebrafish
401	36	78.3 168462	5	BX005308	BX005308 Zebrafish	474	36	78.3 209756	14	AC153709	AC153709 Bos tauru
402	36	78.3 169206	9	AC123550	AC123550 Mus muscu	C 475	36	78.3 210050	1	AJ414146	AJ414146 Versinia
403	36	78.3 169345	14	AC151874	AC151874 Didelphis	C 476	36	78.3 211512	9	AC122193	AC122193 Mus muscu
404	36	78.3 169491	8	AL160056	AL160056 Human DNA	C 477	36	78.3 211647	14	AC114167	AC114167 Rattus no
405	36	78.3 170199	5	BX323793	BX323793 Zebrafish	C 478	36	78.3 212595	9	AC121799	AC121799 Mus muscu
406	36	78.3 170327	2	AC006467	AC006467 Drosophil	479	36	78.3 212908	9	AC164289	AC164289 Mus muscu
407	36	78.3 170336	9	AC125137	AC125137 Mus muscu	480	36	78.3 213473	9	AC140287	AC140287 Mus muscu
408	36	78.3 171206	8	AC090946	AC090946 Homo sapi	C 481	36	78.3 214923	9	AC158611	AC158611 Mus muscu
409	36	78.3 171922	8	AL928636	AL928636 Human DNA	482	36	78.3 217001	9	AC103419	AC103419 Rattus no
410	36	78.3 172054	14	CR853281	CR853281 Danio rer	483	36	78.3 217813	9	AC138739	AC138739 Mus muscu
411	36	78.3 173910	8	AC083982	AC083982 Homo sapi	484	36	78.3 218529	14	CR854851	CR854851 Bos tauru
412	36	78.3 174070	8	AC146092	AC146092 Pan trogl	485	36	78.3 218880	14	AC150974	AC150974 Danio rer
413	36	78.3 174086	14	AC110010	AC110010 Homo sapi	486	36	78.3 219730	9	AC087556	AC087556 Mus muscu
414	36	78.3 174616	9	AC109307	AC109307 Mus muscu	C 487	36	78.3 220746	9	AC161114	AC161114 Mus muscu
415	36	78.3 174912	8	AC103586	AC103586 Homo sapi	C 488	36	78.3 220851	5	BX649335	BX649335 Zebrafish
416	36	78.3 175123	9	AC157275	AC157275 Mus muscu	489	36	78.3 220918	9	AC166574	AC166574 Mus muscu
417	36	78.3 175931	8	H973B6	AL163193 Homo sapi	490	36	78.3 221585	14	AC097412	AC097412 Rattus no
418	36	78.3 176027	14	AL732575	AL732575 Mus muscu	491	36	78.3 221716	14	AC150394	AC150394 Branchios
419	36	78.3 177633	9	AC159102	AC159102 Mus muscu	C 492	36	78.3 221859	9	AL732309	AL732309 Mouse DNA
420	36	78.3 177795	8	AC067743	AC067743 Homo sapi	493	36	78.3 222585	14	AC106558	AC106558 Rattus no
421	36	78.3 178051	14	AC141333	AC141333 Rattus no	494	36	78.3 222698	9	AC140054	AC140054 Mus muscu
422	36	78.3 178416	14	BX936329	BX936329 Danio rer	C 495	36	78.3 222856	14	AC128437	AC128437 Rattus no
423	36	78.3 178929	5	BX470133	BX470133 Zebrafish	496	36	78.3 223041	14	AC144469	AC144469 Canis fam
424	36	78.3 179233	14	CR925824	CR925824 Danio rer	C 497	36	78.3 223171	14	AC162608	AC162608 Bos tauru
425	36	78.3 179236	14	AC022074	AC022074 Homo sapi	C 498	36	78.3 223675	8	AC025031	AC025031 Homo sapi
426	36	78.3 180126	14	AC146189	AC146189 Pan trogl	C 499	36	78.3 224003	9	AC140326	AC140326 Mus muscu
427	36	78.3 180597	8	CR847540	CR847540 Danio rer	C 500	36	78.3 224059	14	AC112742	AC112742 Rattus no
428	36	78.3 180665	14	AC139712	AC139712 Homo sapi	501	36	78.3 224076	14	AC119090	AC119090 Rattus no
429	36	78.3 180898	8	AC116770	AC116770 Mus muscu	C 502	36	78.3 224076	14	AC119090	AC119090 Rattus no
430	36	78.3 183783	5	BX539325	BX539325 Zebrafish	C 503	36	78.3 225538	9	AC100600	AC100600 Mus muscu
431	36	78.3 184284	14	AC162436	AC162436 Lemur cat	C 504	36	78.3 228950	14	AC111512	AC111512 Rattus no
432	36	78.3 184821	9	AC159636	AC159636 Mus muscu	C 505	36	78.3 229085	9	AC142112	AC142112 Mus muscu
433	36	78.3 185624	8	AC108162	AC108162 Homo sapi	C 506	36	78.3 229752	9	AC102652	AC102652 Mus muscu
434	36	78.3 185666	14	AC101747	AC101747 Mus muscu	C 507	36	78.3 229843	14	AC133306	AC133306 Rattus no
435	36	78.3 185926	5	BX548066	BX548066 Zebrafish	C 508	36	78.3 230525	14	AC096982	AC096982 Rattus no
436	36	78.3 186495	14	AC132965	AC132965 Rattus no	509	36	78.3 231679	14	AC123319	AC123319 Rattus no
437	36	78.3 186504	8	AC006040	AC006040 Homo sapi	C 510	36	78.3 231804	14	AC152688	AC152688 Bos tauru
438	36	78.3 187746	14	AC087123	AC087123 Mus muscu	C 511	36	78.3 232065	14	AC069425	AC069425 Homo sapi
439	36	78.3 188192	8	AL139241	AL139241 Human DNA	512	36	78.3 232433	14	AC154192	AC154192 Homo sapi
440	36	78.3 189025	14	AC128007	AC128007 Rattus no	513	36	78.3 233349	14	AC117203	AC117203 Mus muscu
441	36	78.3 190196	5	CR293524	CR293524 Zebrafish	514	36	78.3 233976	14	AC118078	AC118078 Rattus no

c 515	36	78.3	235010	14	AC099447	AC099447 Rattus no	588	35	76.1	582	8	AB209165	AB209165 Homo sapi
c 516	36	78.3	240778	14	AC105584	AC105584 Rattus no	589	35	76.1	583	8	BC070235	BC070235 Homo sapi
c 517	36	78.3	240997	9	AC149091	AC149091 Mus muscu	590	35	76.1	585	6	BD034652	BD034652 Sequence
c 518	36	78.3	243648	14	AC128576	AC128576 Rattus no	591	35	76.1	585	6	AX899119	AX899119 Sequence
c 519	36	78.3	244146	14	AC126704	AC126704 Rattus no	592	35	76.1	585	6	BC001032	BC001032 Homo sapi
c 520	36	78.3	246376	14	AC118353	AC118353 Rattus no	593	35	76.1	597	8	BC071946	BC071946 Homo sapi
c 521	36	78.3	247594	14	AC106655	AC106655 Rattus no	594	35	76.1	598	6	CS032885	CS032885 Sequence
c 522	36	78.3	247779	14	CR925731	CR925731 Danio rer	595	35	76.1	598	6	CS041837	CS041837 Sequence
c 523	36	78.3	247818	14	AC151857	AC151857 Papio ham	596	35	76.1	601	10	BV064131	BV064131 S212P6320
c 524	36	78.3	248214	14	AL138811	AL138811 Homo sapi	597	35	76.1	629	8	BC071991	BC071991 Homo sapi
c 525	36	78.3	248378	14	AC128989	AC128989 Rattus no	598	35	76.1	657	8	BC001955	BC001955 Homo sapi
c 526	36	78.3	251928	14	AC164386	AC164386 Bos tauru	599	35	76.1	657	8	BC005012	BC005012 Homo sapi
c 527	36	78.3	252440	5	BX005425	BX005425 Zebrafish	600	35	76.1	664	10	BV053710	BV053710 S212P6002
c 528	36	78.3	255881	14	AC099252	AC099252 Rattus no	601	35	76.1	672	8	BC073799	BC073799 Homo sapi
c 529	36	78.3	256548	14	AC132542	AC132542 Rattus no	602	35	76.1	720	15	CNS01B89	AL114625 Botrytis
c 530	36	78.3	258754	14	AC128436	AC128436 Rattus no	603	35	76.1	732	10	BV596144	BV596144 S215P6009
c 531	36	78.3	259704	14	AC112092	AC112092 Rattus no	604	35	76.1	751	10	BV663598	BV663598 S215P6781
c 532	36	78.3	259933	14	AC106643	AC106643 Rattus no	605	35	76.1	763	10	BV625363	BV625363 S217P6964
c 533	36	78.3	260215	14	AC095852	AC095852 Rattus no	c 606	35	76.1	809	10	BV457548	BV457548 gfh46a10
c 534	36	78.3	264638	14	AC129847	AC129847 Rattus no	607	35	76.1	828	10	BV485632	BV485632 S215P6099
c 535	36	78.3	265132	14	AC153017	AC153017 Mus muscu	608	35	76.1	893	10	BV467240	BV467240 Danio rer
c 536	36	78.3	265185	14	AC117637	AC117637 Mus muscu	c 609	35	76.1	1076	5	BC093270	BC093270 G591P6805
c 537	36	78.3	266143	14	AC165418	AC165418 Mus muscu	c 610	35	76.1	1104	8	AK057942	AK057942 Homo sapi
c 538	36	78.3	267809	14	CR626865	CR626865 Danio rer	611	35	76.1	1199	6	CQ714092	CQ714092 Sequence
c 539	36	78.3	268099	14	AC133407	AC133407 Rattus no	612	35	76.1	1248	8	BC000393	BC000393 Homo sapi
c 540	36	78.3	268505	14	AC152864	AC152864 Pan trogl	613	35	76.1	1248	8	BC019300	BC019300 Homo sapi
c 541	36	78.3	268651	14	BX005198	BX005198 Danio rer	614	35	76.1	1281	6	BD249732	BD249732 Polypteri
c 542	36	78.3	268694	14	AC095697	AC095697 Rattus no	615	35	76.1	1281	6	AX025012	AX025012 Sequence
c 543	36	78.3	269040	14	AC099002	AC099002 Rattus no	616	35	76.1	1281	6	AX316087	AX316087 Sequence
c 544	36	78.3	269542	14	AC121618	AC121618 Rattus no	617	35	76.1	1331	3	AV592115	AV592115 Unculture
c 545	36	78.3	271150	14	AC162407	AC162407 Bos tauru	618	35	76.1	1331	9	AF548356	AF548356 Cricetulu
c 546	36	78.3	271646	14	AC123484	AC123484 Rattus no	619	35	76.1	1573	6	AF577875	AF577875 Sequence
c 547	36	78.3	272174	5	CR388032	CR388032 Zebrafish	c 620	35	76.1	1626	15	AF082890	AF082890 Solanum t
c 548	36	78.3	272247	14	AC115350	AC115350 Rattus no	621	35	76.1	1766	6	CQ843410	CQ843410 Sequence
c 549	36	78.3	276137	14	AC131170	AC131170 Rattus no	622	35	76.1	1766	8	AK124325	AK124325 Homo sapi
c 550	36	78.3	277603	14	AC079543	AC079543 Mus muscu	c 623	35	76.1	1774	6	AK714164	AK714164 Sequence
c 551	36	78.3	288992	14	AC163835	AC163835 Bos tauru	c 624	35	76.1	1774	6	AK056323	AK056323 Homo sapi
c 552	36	78.3	290393	1	AB017140	AB017140 Yersinia	625	35	76.1	2209	9	F289804810	AF289813 Mus muscu
c 553	36	78.3	294258	14	AC156066	AC156066 Bos tauru	626	35	76.1	2333	9	AF063939	AF289813 Rattus no
c 554	36	78.3	298116	2	AB003782	AB003782 Drosophila	627	35	76.1	2361	9	BC087011	BC087011 Rattus no
c 555	36	78.3	306309	14	AC105494	AC105494 Rattus no	628	35	76.1	2421	6	CQ903899	CQ903899 Sequence
c 556	36	78.3	328128	2	AB003486	AB003486 Drosophila	629	35	76.1	2421	9	AF314636	AF314636 Mus muscu
c 557	35	76.1	166	6	AX396059	AX396059 Sequence	630	35	76.1	2423	9	BC058198	BC058198 Mus muscu
c 558	35	76.1	168	6	AX240262	AX240262 Sequence	631	35	76.1	2557	6	AX961912	AX961912 Sequence
c 559	35	76.1	174	8	HSBPF42	270761 H. sapiens m	632	35	76.1	2557	6	AX961914	AX961914 Sequence
c 560	35	76.1	201	10	BV201889	BV201889 sqm20764	633	35	76.1	2567	9	AF316985	AF316985 Mus muscu
c 561	35	76.1	245	6	CQ665058	CQ665058 Sequence	634	35	76.1	2567	9	AY009154	AY009154 Mus muscu
c 562	35	76.1	253	6	CQ659192	CQ659192 Sequence	635	35	76.1	2604	6	CQ903897	CQ903897 Sequence
c 563	35	76.1	257	6	CQ661964	CQ661964 Sequence	636	35	76.1	2604	6	CQ903898	CQ903898 Sequence
c 564	35	76.1	277	6	CQ666246	CQ666246 Sequence	637	35	76.1	2604	6	E34465	E34465 Novel Toll-
c 565	35	76.1	321	6	CQ668645	CQ668645 Sequence	638	35	76.1	2604	9	AB020808	AB020808 Mus muscu
c 566	35	76.1	344	6	CQ6201917	CQ6201917 Sequence	639	35	76.1	2625	9	BC055366	BC055366 Mus muscu
c 567	35	76.1	383	6	CQ662298	CQ662298 Sequence	c 640	35	76.1	2625	9	AX835149	AX835149 Sequence
c 568	35	76.1	398	6	AX245109	AX245109 Sequence	c 641	35	76.1	2642	8	AK098006	AK098006 Homo sapi
c 569	35	76.1	406	6	AX773867	AX773867 Sequence	642	35	76.1	2713	9	BC081791	BC081791 Rattus no
c 570	35	76.1	407	2	PFAS36712	PFAS36712 Plasmodi	643	35	76.1	2715	6	CQ605235	CQ605235 Sequence
c 571	35	76.1	425	6	CQ919730	CQ919730 Sequence	644	35	76.1	2820	9	BC062390	BC062390 Rattus no
c 572	35	76.1	456	10	G97781	G97781 S208P6505PC	c 645	35	76.1	2854	1	TENMOAB	TENMOAB Bacteroides
c 573	35	76.1	461	10	G06829	G06829 human STS W	646	35	76.1	2934	5	AJ720279	AJ720279 Gallus ga
c 574	35	76.1	475	6	AX360373	AX360373 Sequence	647	35	76.1	2997	9	MMU14390	MMU14390 Mus muscu
c 575	35	76.1	498	11	AY889711	AY889711 Synthetic	648	35	76.1	3194	8	AK025353	AK025353 Homo sapi
c 576	35	76.1	498	11	AY892166	AY892166 Synthetic	649	35	76.1	3238	4	BC069041	BC069041 Mus muscu
c 577	35	76.1	502	10	BV069451	BV069451 S212P6073	650	35	76.1	3240	4	AY525124	AY525124 Canis fam
c 578	35	76.1	537	6	BD204669	BD204669 Human nuc	651	35	76.1	3349	9	DA9802	DA9802 Mus muscu
c 579	35	76.1	537	6	AX014099	AX014099 Sequence	c 652	35	76.1	3655	5	CR942651	CR942651 Xenopus t
c 580	35	76.1	558	10	BV271014	BV271014 S235P6190	c 653	35	76.1	3997	2	BT010317	BT010317 Drosophila
c 581	35	76.1	559	6	CQ715199	CQ715199 Sequence	c 654	35	76.1	4172	8	CQ850206	CQ850206 Sequence
c 582	35	76.1	559	6	CQ723277	CQ723277 Sequence	c 655	35	76.1	4172	8	AK127317	AK127317 Homo sapi
c 583	35	76.1	562	6	CQ728718	CQ728718 Sequence	c 656	35	76.1	4435	6	CS124634	CS124634 Sequence
c 584	35	76.1	566	6	CQ715198	CQ715198 Sequence	c 657	35	76.1	4761	8	HSM07187	HSM07187 Homo sapien
c 585	35	76.1	570	6	AX410660	AX410660 Sequence	c 658	35	76.1	5240	8	BX641132	BX641132 Homo sapi
c 586	35	76.1	570	8	HSU14972	U14972 Human ribos	659	35	76.1	5893	6	CQ605234	CQ605234 Sequence
c 587	35	76.1	572	6	CQ723492	CQ723492 Sequence	660	35	76.1	5965	9	BC072574	BC072574 Mus muscu

661	35	76.1	6234	6	AR585216	AR585216 Sequence	734	35	76.1	104614	5	BX571758	BX571758 Zebrafish
662	35	76.1	6629	9	MMU56439	U56439 Mus musculus	735	35	76.1	104648	14	BX682534	BX682534 Homo sapi
c 663	35	76.1	7122	6	AR144957	AR144957 Sequence	c 736	35	76.1	104806	15	AC006228	AC006228 Genomic s
c 664	35	76.1	7122	6	AR367907	AR367907 Sequence	c 737	35	76.1	106019	14	AP007431	AP007431 Lotus cor
c 665	35	76.1	7122	6	AX050442	AX050442 Sequence	c 738	35	76.1	108975	15	CR762430	CR762430 Zebrafish
c 666	35	76.1	7122	6	AX069340	AX069340 Sequence	c 739	35	76.1	110000	1	AE008692_05	Continuation (6 of
c 667	35	76.1	7122	6	HSU75743	U75743 Human methi	c 740	35	76.1	110000	1	AE017282_11	Continuation (12 o
c 668	35	76.1	7224	6	AR300095	AR300095 Sequence	741	35	76.1	110000	1	BA000283_32	Continuation (33 o
c 669	35	76.1	7224	6	AR367908	AR367908 Sequence	742	35	76.1	110000	1	BA000036_20	Continuation (21 o
c 670	35	76.1	7224	6	AR438495	AR438495 Sequence	c 743	35	76.1	110000	1	EX571965_08	Continuation (9 of
c 671	35	76.1	7224	6	AR562003	AR562003 Sequence	c 744	35	76.1	110000	1	CP000082_09	Continuation (10 o
c 672	35	76.1	7224	6	AX069342	AX069342 Sequence	745	35	76.1	110000	1	CP000084_00	CP000084 Candidatu
c 673	35	76.1	7224	6	HSU73338	U73338 Human methi	746	35	76.1	110000	14	AC106698_1	Continuation (2 of
c 674	35	76.1	7224	6	AX251111	AX251111 Sequence	747	35	76.1	110000	14	BX890561_0	BX890561 Danio rer
c 675	35	76.1	7942	6	MMU012160	AJ012160 Mus muscu	748	35	76.1	110000	14	CR382381_2	Continuation (3 of
c 676	35	76.1	9115	1	AB024693	AB024693 Leptospir	c 749	35	76.1	110000	15	CR380953_08	Continuation (9 of
c 677	35	76.1	10455	1	AB011496	AB011496 Leptospir	c 750	35	76.1	110000	15	CR382135_05	Continuation (11 o
c 678	35	76.1	11332	1	AB013471	AB013471 Methanosa	c 751	35	76.1	110000	15	AB017345_10	Continuation (11 o
c 679	35	76.1	11483	1	AE003930	AE003930 Xylella f	752	35	76.1	110791	5	EX784024	EX784024 Zebrafish
c 680	35	76.1	12559	8	AL772134	AL772134 Human DNA	753	35	76.1	111678	15	AC149575	AC149575 Populus t
c 681	35	76.1	17815	8	AL513010	AL513010 Human DNA	c 754	35	76.1	111775	9	AC018559	AC018559 Mus muscu
c 682	35	76.1	24300	13	CHB131832	AJ131832 Gallid he	c 755	35	76.1	112467	8	HS058B13	Z98052 Human DNA s
c 683	35	76.1	28607	2	U23179	U23179 Caenorhabdi	c 756	35	76.1	115069	15	AP006071	AP006071 Lotus cor
c 684	35	76.1	34013	14	AC010013	AC10013 Drosophil	757	35	76.1	115968	8	AL139100	AL139100 Human DNA
685	35	76.1	34038	8	AC153808	AC153808 Homo sapi	758	35	76.1	116299	14	RN141N5	AL603725 Rattus no
686	35	76.1	42591	6	BD129559	BD129559 Polynucle	c 759	35	76.1	116452	14	AC160593	AC160593 Loxodonta
687	35	76.1	42522	2	U64857	U64857 Caenorhabdi	c 760	35	76.1	116601	15	AC149634	AC149634 Medicago
c 688	35	76.1	48915	5	BX323554	BX323554 Zebrafish	761	35	76.1	119203	14	AC138155	AC138155 Carollia
c 689	35	76.1	49377	8	AL391316	AL391316 Human DNA	762	35	76.1	119745	15	AC141108	AC141108 Medicago
c 690	35	76.1	57693	14	AC110044	AC110044 Homo sapi	c 763	35	76.1	119806	8	AL137795	AL137795 Human DNA
c 691	35	76.1	57904	14	AC091648	AC091648 Homo sapi	c 764	35	76.1	120017	15	AC149572	AC149572 Mus muscu
c 692	35	76.1	58037	5	AL713396	AL713396 Zebrafish	c 765	35	76.1	121140	14	AC166371	AC166371 Homo sapi
c 693	35	76.1	58253	14	CR382381_3	Continuation (4 of	c 766	35	76.1	121743	14	CT009763	CT009763 Danio rer
c 694	35	76.1	59157	14	AC165678	AC165678 Bos tauru	c 767	35	76.1	122568	8	AF196972	AF196972 Homo sapi
c 695	35	76.1	61502	14	AC090765	AC090765 Homo sapi	768	35	76.1	122915	15	AC157984	AC157984 Medicago
c 696	35	76.1	63308	8	AC112490	AC112490 Homo sapi	c 769	35	76.1	124524	8	EX255972	EX255972 Human DNA
c 697	35	76.1	64079	14	AC104585	AC104585 Homo sapi	770	35	76.1	125123	8	AC104455	AC104455 Homo sapi
c 698	35	76.1	65141	5	BX294107	BX294107 Zebrafish	c 771	35	76.1	126259	8	BSA371L19	AL118502 Human DNA
c 699	35	76.1	66694	5	BX908791	BX908791 Zebrafish	c 772	35	76.1	127718	8	HS000059	BS000059 Pan trogl
c 700	35	76.1	67201	8	AL136169	AL136169 Human DNA	773	35	76.1	131853	14	AC017755	AC017755 Drosophil
c 701	35	76.1	67709	8	AC104830	AC104830 Homo sapi	c 774	35	76.1	132983	14	AC136591	AC136591 Homo sapi
c 702	35	76.1	68007	14	AC152929	AC152929 Xenopus t	c 775	35	76.1	133130	14	AC141920	AC141920 Rattus no
c 703	35	76.1	68041	15	AC025814	AC025814 Arabidops	c 776	35	76.1	133526	14	AC105371	AC105371 Sus scrofa
c 704	35	76.1	68402	14	AC100370	AC100370 Mus muscu	777	35	76.1	134095	14	AC006915	AC006915 Caenorhab
c 705	35	76.1	69217	14	AC129516	AC129516 Homo sapi	c 778	35	76.1	135380	4	CR753179	CR753179 Opossum D
c 706	35	76.1	69784	14	AC090746	AC090746 Homo sapi	c 779	35	76.1	135468	14	AC016503	AC016503 Homo sapi
c 707	35	76.1	70106	8	AC024293	AC024293 Homo sapi	c 780	35	76.1	136371	8	AC025451	AC025451 Zebrafish
c 708	35	76.1	72245	8	HS93C23	AL109836 Human DNA	c 781	35	76.1	138235	5	CR788281	CR788281 Zebrafish
c 709	35	76.1	72594	8	HSBA243C2	Continuation (4 of	c 782	35	76.1	138264	9	AC132579	AC132579 Mus muscu
c 710	35	76.1	79834	14	AC157811_3	AC157712 Xenopus t	783	35	76.1	138829	5	EX571762	EX571762 Zebrafish
c 711	35	76.1	79885	14	AC157712	AL591180 Zebrafish	784	35	76.1	138973	9	AC090712	AC090712 Mus muscu
c 712	35	76.1	80272	5	AL591180	CR382331 Zebrafish	c 785	35	76.1	138994	14	AC141046	AC141046 Rattus no
c 713	35	76.1	80921	5	CR382331	CR382331 Zebrafish	c 786	35	76.1	139530	5	EX545917	EX545917 Zebrafish
c 714	35	76.1	82014	14	CR376841	CR376841 Danio rer	787	35	76.1	139594	14	AC025085	AC025085 Homo sapi
c 715	35	76.1	82906	8	HS970A17	AL034431 Human DNA	c 788	35	76.1	140801	14	AC015663	AC015663 Homo sapi
c 716	35	76.1	83763	5	AL935140	AL935140 Zebrafish	c 789	35	76.1	141407	14	EX936452	EX936452 Danio rer
c 717	35	76.1	85002	14	AC156003	AC156003 Xenopus t	c 790	35	76.1	141415	14	AC008032	AC008032 Homo sapi
c 718	35	76.1	86871	8	AC106755	AC106755 Homo sapi	c 791	35	76.1	142064	14	EX571896	EX571896 Danio rer
c 719	35	76.1	86954	8	AC026430	AC026430 Homo sapi	c 792	35	76.1	142063	8	AC091193	AC091193 Homo sapi
c 720	35	76.1	92211	15	CR380953_09	Continuation (10 o	c 793	35	76.1	142488	9	AC087040	AC087040 Mus muscu
c 721	35	76.1	92873	14	AC165655	AC165655 Bos tauru	794	35	76.1	142908	14	AL513172	AL513172 Homo sapi
c 722	35	76.1	95908	8	AC008628	AC008628 Homo sapi	795	35	76.1	143486	5	EX571830	EX571830 Zebrafish
c 723	35	76.1	96546	5	AC151818	AC151818 Xenopus t	796	35	76.1	144869	14	AC144884	AC144884 Rattus no
c 724	35	76.1	96688	14	AC153685	AC153685 Bos tauru	797	35	76.1	145626	14	AC020721	AC020721 Homo sapi
c 725	35	76.1	97351	8	AC015853	AC015853 Homo sapi	798	35	76.1	146065	5	EX276112	EX276112 Zebrafish
c 726	35	76.1	99108	5	EX537146	EX537146 Zebrafish	799	35	76.1	146572	9	AC118259	AC118259 Mus muscu
c 727	35	76.1	99866	5	AL591442	AL591442 Zebrafish	800	35	76.1	146786	5	AL772279	AL772279 Zebrafish
c 728	35	76.1	100000	8	AP000056	AP000056 Homo sapi	801	35	76.1	146810	8	AL159170	AL159170 Human DNA
c 729	35	76.1	100000	8	AP000171	AP000171 Homo sapi	802	35	76.1	146850	14	CR394559	CR394559 Danio rer
c 730	35	76.1	100510	5	EX571978	EX571978 Zebrafish	803	35	76.1	146978	8	AL160272	AL160272 Human DNA
c 731	35	76.1	100859	5	AL929128	AL929128 Zebrafish	c 804	35	76.1	147474	14	AL844173	AL844173 Homo sapi
c 732	35	76.1	101015	5	EX649466	EX649466 Zebrafish	c 805	35	76.1	147708	8	HS179M20	Z97053 Human DNA s
c 733	35	76.1	102243	14	AC153091	AC153091 Ornithorh	c 806	35	76.1	148116	14	AC078838	AC078838 Homo sapi

807	35	76.1	148379	9	AC139321	Mus muscu	AC139321	Mus muscu	AC114563	9	AC114563	AC114563	Mus muscu
c 808	35	76.1	148965	9	AC107692	Mus muscu	AC107692	Mus muscu	AL954741	5	AL954741	AL954741	Zebrafish
809	35	76.1	149684	9	AC154669	Mus muscu	AC154669	Mus muscu	EX005359	14	EX005359	EX005359	Danio rer
c 810	35	76.1	149817	14	AC069440	Homo sapi	AC069440	Homo sapi	AC158516	9	AC158516	AC158516	Mus muscu
811	35	76.1	150195	14	CR847968	Danio rer	CR847968	Danio rer	AC146927	14	AC146927	AC146927	Atelalex
c 812	35	76.1	150600	14	CR847828	Danio rer	CR847828	Danio rer	AY016023	5	AY016023	AY016023	Sphoeroid
813	35	76.1	150660	14	AL157372	Human DNA	AL157372	Human DNA	AC124428	9	AC124428	AC124428	Mus muscu
c 814	35	76.1	150719	14	CR450718	Danio rer	CR450718	Danio rer	AC073832	14	AC073832	AC073832	Homo sapi
c 815	35	76.1	151017	8	AL591431	Human DNA	AL591431	Human DNA	AC091911	8	AC091911	AC091911	Homo sapi
816	35	76.1	151386	14	AC013312	Homo sapi	AC013312	Homo sapi	AC144408	8	AC144408	AC144408	Mus muscu
c 817	35	76.1	151517	8	AC105230	Homo sapi	AC105230	Homo sapi	EX957242	5	EX957242	EX957242	Zebrafish
818	35	76.1	151605	8	AC011995	Homo sapi	AC011995	Homo sapi	AC084260	8	AC084260	AC084260	Homo sapi
c 819	35	76.1	151904	14	EX908739	Danio rer	EX908739	Danio rer	AC092472	8	AC092472	AC092472	Homo sapi
c 820	35	76.1	152316	9	AC007844	Mus muscu	AC007844	Mus muscu	CR19474	14	CR19474	CR19474	Rattus no
c 821	35	76.1	152321	14	AC136935	Homo sapi	AC136935	Homo sapi	CR792455	14	CR792455	CR792455	Danio rer
822	35	76.1	152617	5	EX005387	Zebrafish	EX005387	Zebrafish	EX936450	5	EX936450	EX936450	Zebrafish
c 823	35	76.1	152698	14	AC140802	Mus muscu	AC140802	Mus muscu	AC016135	8	AC016135	AC016135	Homo sapi
c 824	35	76.1	153127	5	EX004870	Zebrafish	EX004870	Zebrafish	EX005378	5	EX005378	EX005378	Zebrafish
825	35	76.1	153612	14	EX004806	Danio rer	EX004806	Danio rer	AC093121	2	AC093121	AC093121	Drosophila
c 826	35	76.1	153759	14	CR848039	Danio rer	CR848039	Danio rer	AL390024	14	AL390024	AL390024	Homo sapi
c 827	35	76.1	153989	5	AC005074	Zebrafish	AC005074	Zebrafish	EX005081	5	EX005081	EX005081	Zebrafish
c 828	35	76.1	154256	5	EX511192	Zebrafish	EX511192	Zebrafish	EX470264	5	EX470264	EX470264	Zebrafish
829	35	76.1	154732	14	AC023236	Homo sapi	AC023236	Homo sapi	AC023985	14	AC023985	AC023985	Homo sapi
c 830	35	76.1	155237	5	CR855305	Zebrafish	CR855305	Zebrafish	CR847948	5	CR847948	CR847948	Zebrafish
c 831	35	76.1	155340	8	AC079856	Homo sapi	AC079856	Homo sapi	AC034193	8	AC034193	AC034193	Homo sapi
832	35	76.1	155348	5	AL928820	Zebrafish	AL928820	Zebrafish	AC073460	5	AC073460	AC073460	Homo sapi
c 833	35	76.1	155920	5	EX005182	Zebrafish	EX005182	Zebrafish	AC087865	5	AC087865	AC087865	Homo sapi
834	35	76.1	156276	14	CR1769765	Danio rer	CR1769765	Danio rer	AL928724	14	AL928724	AL928724	Danio rer
c 835	35	76.1	156897	14	AC150303	Papio anu	AC150303	Papio anu	AC144566	14	AC144566	AC144566	Homo sapi
c 836	35	76.1	157052	5	AL929459	Zebrafish	AL929459	Zebrafish	AC060766	5	AC060766	AC060766	Homo sapi
c 837	35	76.1	157052	14	AC090545	Homo sapi	AC090545	Homo sapi	AC137873	14	AC137873	AC137873	Homo sapi
838	35	76.1	157330	5	AL808129	Zebrafish	AL808129	Zebrafish	AL935189	5	AL935189	AL935189	Zebrafish
c 839	35	76.1	157362	14	EX004848	Danio rer	EX004848	Danio rer	AC078896	14	AC078896	AC078896	Mus muscu
840	35	76.1	157412	5	EX513790	Zebrafish	EX513790	Zebrafish	AC154853	5	AC154853	AC154853	Mus muscu
c 841	35	76.1	157422	14	CR405707	Danio rer	CR405707	Danio rer	AP000330	14	AP000330	AP000330	Homo sapi
c 842	35	76.1	157508	8	AC018422	Homo sapi	AC018422	Homo sapi	EX005481	8	EX005481	EX005481	Zebrafish
c 843	35	76.1	157921	5	EX950217	Zebrafish	EX950217	Zebrafish	EX890557	5	EX890557	EX890557	Zebrafish
844	35	76.1	158660	14	AC068425	Mus muscu	AC068425	Mus muscu	AL935131	14	AL935131	AL935131	Zebrafish
c 845	35	76.1	158696	9	AC133581	Mus muscu	AC133581	Mus muscu	AC090097	9	AC090097	AC090097	Homo sapi
c 846	35	76.1	158998	8	AC091516	Homo sapi	AC091516	Homo sapi	EX649397	8	EX649397	EX649397	Zebrafish
c 847	35	76.1	159621	8	AC083801	Homo sapi	AC083801	Homo sapi	AC150136	8	AC150136	AC150136	Gallus ga
c 848	35	76.1	159878	14	CR956623	Danio rer	CR956623	Danio rer	AC124708	14	AC124708	AC124708	Mus muscu
c 849	35	76.1	159930	14	AC118047	Mus muscu	AC118047	Mus muscu	EX323998	14	EX323998	EX323998	Zebrafish
850	35	76.1	159984	14	CR394451	Danio rer	CR394451	Danio rer	AC087170	14	AC087170	AC087170	Homo sapi
c 851	35	76.1	160029	5	EX000984	Zebrafish	EX000984	Zebrafish	AL157398	5	AL157398	AL157398	Human DNA
c 852	35	76.1	160247	5	EX927247	Zebrafish	EX927247	Zebrafish	AL159259	5	AL159259	AL159259	Human DNA
c 853	35	76.1	160456	8	AC019011	Homo sapi	AC019011	Homo sapi	AC142168	8	AC142168	AC142168	Mus muscu
c 854	35	76.1	160555	5	EX296516	Zebrafish	EX296516	Zebrafish	AC102195	5	AC102195	AC102195	Mus muscu
c 855	35	76.1	160631	5	CR450753	Zebrafish	CR450753	Zebrafish	AP007207	5	AP007207	AP007207	Mus muscu
856	35	76.1	160738	14	AC161787	Cercopithec	AC161787	Cercopithec	AC155839	14	AC155839	AC155839	Mus muscu
c 857	35	76.1	160775	9	AC125158	Mus muscu	AC125158	Mus muscu	AC093501	9	AC093501	AC093501	Drosophila
c 858	35	76.1	161002	14	CR759944	Danio rer	CR759944	Danio rer	AP002356	14	AP002356	AP002356	Homo sapi
c 859	35	76.1	161378	14	AC136633	Homo sapi	AC136633	Homo sapi	AC154761	14	AC154761	AC154761	Mus muscu
c 860	35	76.1	161602	8	AP001978	Homo sapi	AP001978	Homo sapi	EX323875	8	EX323875	EX323875	Zebrafish
861	35	76.1	161671	14	AC018369	Homo sapi	AC018369	Homo sapi	AC006169	14	AC006169	AC006169	Drosophila
c 862	35	76.1	161731	8	AC107021	Homo sapi	AC107021	Homo sapi	CR407551	8	CR407551	CR407551	Danio rer
863	35	76.1	161737	14	AC087809	Homo sapi	AC087809	Homo sapi	EX511240	14	EX511240	EX511240	Zebrafish
c 864	35	76.1	161935	8	AC008269	Homo sapi	AC008269	Homo sapi	EX547997	8	EX547997	EX547997	Zebrafish
c 865	35	76.1	162474	8	AC022706	Homo sapi	AC022706	Homo sapi	AC011912	8	AC011912	AC011912	Homo sapi
c 866	35	76.1	162515	8	AC091097	Homo sapi	AC091097	Homo sapi	AL445258	8	AL445258	AL445258	Human DNA
c 867	35	76.1	162858	9	AC132598	Mus muscu	AC132598	Mus muscu	AC012108	9	AC012108	AC012108	Homo sapi
868	35	76.1	163452	14	AC079184	Homo sapi	AC079184	Homo sapi	CR407558	14	CR407558	CR407558	Zebrafish
c 869	35	76.1	163525	5	EX005048	Zebrafish	EX005048	Zebrafish	CR759777	5	CR759777	CR759777	Danio rer
c 870	35	76.1	163791	14	AC034135	Homo sapi	AC034135	Homo sapi	AC155645	14	AC155645	AC155645	Mus muscu
c 871	35	76.1	164499	5	CR407585	Zebrafish	CR407585	Zebrafish	AC080149	5	AC080149	AC080149	Homo sapi
872	35	76.1	164529	14	AC079975	Homo sapi	AC079975	Homo sapi	AC152452	14	AC152452	AC152452	Mus muscu
c 873	35	76.1	164741	5	EX005455	Zebrafish	EX005455	Zebrafish	EX470137	5	EX470137	EX470137	Danio rer
c 874	35	76.1	164897	14	AC015664	Homo sapi	AC015664	Homo sapi	CR450794	14	CR450794	CR450794	Danio rer
c 875	35	76.1	164933	14	AC074231	Homo sapi	AC074231	Homo sapi	AC122839	14	AC122839	AC122839	Mus muscu
876	35	76.1	164950	5	CR788314	Zebrafish	CR788314	Zebrafish	AC145438	5	AC145438	AC145438	Homo sapi
c 877	35	76.1	165311	8	AC097061	Homo sapi	AC097061	Homo sapi	CR354422	8	CR354422	CR354422	Zebrafish
c 878	35	76.1	165973	14	AC159648	Mus muscu	AC159648	Mus muscu	AC113687	14	AC113687	AC113687	Rattus no
c 879	35	76.1	166119	14	EX950197	Danio rer	EX950197	Danio rer					

```

c 953 35 76.1 178225 5 BX005286
c 954 35 76.1 178230 8 AC018828
c 955 35 76.1 178500 4 AL359926
c 956 35 76.1 179095 5 AL954704
c 957 35 76.1 179165 14 AC140289
c 958 35 76.1 179292 14 AC022632
c 959 35 76.1 179396 14 BX649297
c 960 35 76.1 179396 14 AC017056
c 961 35 76.1 179414 14 AC165320
c 962 35 76.1 179576 1 BX234156
c 963 35 76.1 179741 14 AC116815
c 964 35 76.1 180042 14 CR388184
c 965 35 76.1 180353 8 AC079326
c 966 35 76.1 180483 5 BX469901
c 967 35 76.1 180816 14 AC110395
c 968 35 76.1 180861 14 CR391970
c 969 35 76.1 181133 14 AC073478
c 970 35 76.1 181168 5 BX293540
c 971 35 76.1 181942 9 AC165145
c 972 35 76.1 182033 5 BX323544
c 973 35 76.1 182223 14 CR790384
c 974 35 76.1 182325 14 AC121193
c 975 35 76.1 182948 9 AC133183
c 976 35 76.1 183229 5 BX465223
c 977 35 76.1 183298 14 AC025076
c 978 35 76.1 183374 14 AC166098
c 979 35 76.1 183626 14 AC026648
c 980 35 76.1 183804 14 CR854989
c 981 35 76.1 183804 14 CR854989
c 982 35 76.1 183817 14 AC117137
c 983 35 76.1 183847 5 AC145510
c 984 35 76.1 183972 9 AC117223
c 985 35 76.1 184232 5 CR381684
c 986 35 76.1 184254 9 AL773566
c 987 35 76.1 185000 14 AC004064
c 988 35 76.1 185134 8 AC007878
c 989 35 76.1 185388 9 AC122483
c 990 35 76.1 185626 5 BX530068
c 991 35 76.1 185657 14 AC125630
c 992 35 76.1 185849 5 BX005073
c 993 35 76.1 185990 14 CR812836
c 994 35 76.1 186820 5 BX469921
c 995 35 76.1 187795 14 AC025405
c 996 35 76.1 187810 14 CR677513
c 997 35 76.1 188066 14 AC160427
c 998 35 76.1 188631 8 AC022383
c 999 35 76.1 188716 14 AC147323
c1000 35 76.1 189215 9 AC124490

BX005286 Zebrafish
AC018828 Homo sapi
AL359926 Homo sapi
AL954704 Zebrafish
AC140289 Mus muscu
AC022632 Homo sapi
BX649297 Danio rer
AC017056 Homo sapi
AC165320 Mus muscu
BX294156 Pirellula
AC116815 Mus muscu
CR388184 Danio rer
AC079326 Homo sapi
BX469901 Zebrafish
AC110395 Rattus no
CR391970 Danio rer
AC073478 Homo sapi
BX293540 Zebrafish
AC165145 Mus muscu
BX323544 Zebrafish
CR790384 Danio rer
AC121193 Rattus no
AC133183 Mus muscu
BX465223 Zebrafish
AC025076 Homo sapi
AC166098 Mus muscu
AC026648 Homo sapi
CR854989 Danio rer
CR854989 Danio rer
AC117137 Rattus no
AC145510 Gaetost
AC117223 Mus muscu
CR381684 Zebrafish
AL773566 Mouse DNA
AC004064 Homo sapi
AC007878 Homo sapi
AC122483 Mus muscu
BX530068 Zebrafish
AC125630 Rattus no
BX005073 Zebrafish
CR812836 Danio rer
BX469921 Zebrafish
AC025405 Homo sapi
CR677513 Danio rer
AC160427 Bos tauru
AC022383 Homo sapi
AC147323 Pan trogl
AC124490 Mus muscu
```

ALIGNMENTS

```

RESULT 1
LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Mack,D.H., Gish,K.C. and Afar,D.
TITLE Methods for diagnosis of breast cancer, compositions and methods of screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
BOS Biotechnology, Inc. (US)
FEATURES
    source
        1..927
            /organism="Homo sapiens"
```

```

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 3.37 Length: 927
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x AX829164 (1-927)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 241 AATCTGACCGAGGTGCCACGACCTG 267

RESULT 2
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1263)
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1999 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
PI MILRS WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61P35/00,C07K7/06,C07K14/065,
PC C07K19/00,
PC C12N15/00,
CC Polypeptide
FH Key Location/Qualifiers
FT source 1..1263
    /organism="Homo sapiens (human)"
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 4.53 Length: 1263
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x BD249731 (1-1263)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 241 AATCTGACCGAGGTGCCACGACCTG 267

RESULT 3
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
```

```

DEFINITION Sequence 1 from Patent WO029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source Location/Qualifiers
1. .1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.53 Length: 1263
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-11 (1-9) x AX025011 (1-1263)
Qy 1 AnLeuThrGluValProThrAspLeu 9
Db 241 AATCTGACCGAGGTGCCACGGACCTG 267
RESULT 4
AX316086
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source Location/Qualifiers
1. .1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.53 Length: 1263
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-11 (1-9) x AX316086 (1-1263)
Qy 1 AnLeuThrGluValProThrAspLeu 9
Db 241 AATCTGACCGAGGTGCCACGGACCTG 267

```

```

RESULT 5
CQ731678
LOCUS CQ731678 2053 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 17612 from Patent WO02068579.
ACCESSION CQ731678
VERSION CQ731678.1 GI:42308932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 17612 06-SRP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1. .2053
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7.21 Length: 2053
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-11 (1-9) x CQ731678 (1-2053)
Qy 1 AnLeuThrGluValProThrAspLeu 9
Db 325 AATCTGACCGAGGTGCCACGGACCTG 351
RESULT 6
HS5T40A
LOCUS HS5T40A 2053 bp RNA linear PRI 18-APR-2005
DEFINITION Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.
ACCESSION Z29083
VERSION Z29083.1 GI:435654
KEYWORDS 5T4 gene; 5T4 oncofoetal antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Myers,K.A., Rahi-Saund,V., Davison,M.D., Young,J.A., Cheater,A.J.
and Stern,P.L.
TITLE Isolation of a cDNA encoding 5T4 oncofoetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
JOURNAL J. Biol. Chem. 269 (12), 9319-9324 (1994)
PUBMED 8132670
REFERENCE
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
FEATURES
source Location/Qualifiers
1. .2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"

```



```

/misc_RNA      /clone lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label="N-flank"
85..1347
/codon_start=1
/avoidance=experimental
/product="574 oncofetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSS
APFLASAVAPQPPDQCPCSEAAATKVCNRLNTEVPTDLPAYVNLFTGNQ
LAVLPAGAFARRPPLAELALNLSGSLRDEVRAGAFELPSRLDLSHNPLADLSPF
AFSGSNASVSAPSPPLVELILNHI VPPEDSRQNRSGFMVVAALLAGRALQGLRLLELA
SNHFLYLPDVLQALPSLRHLDSLNNLSVSLTYVSNLTHLESLELDNALKVLHNG
TLAEQLGPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPEKMRNVL
LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLLVLYLNKGIKKWMNIRDAC
RDMEGYHYRYEINADPRLNLSNSDV"
130..171
sig_peptide
/misc_RNA      373..966
/product="Leucine rich repeat region"
/label="LRRs"
966..1119
/product="LRR C-terminal flank"
/label="C-flank"
1153..1215
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

ORIGIN
Alignment Scores:
Pred. No.:      7.21      Length:      2053
Score:          46.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:        0

US-10-774-176-11 (1-9) x HS5T40A (1-2053)

QY      1 AsnLeuThrGluValProThrAspLeu 9
|||||
Db      325 AATCTGACCGAGGTGCCACGACCTG 351

RESULT 7
BD127282      2359 bp      DNA      linear      PAT 18-SEP-2002
LOCUS
DEFINITION    Primer for synthesizing full-length cDNA and use thereof.
ACCESSION     BD127282
VERSION       BD127282.1 GI:23222227
KEYWORDS      JP 2002017375-A/2713.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE     1 (bases 1 to 2359)
AUTHORS      Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
              Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
              Koga,H.
              Primers for synthesizing full-length cDNA and use thereof
              Patent: JP 2002017375-A/2713 22-JAN-2002;
              HELIX RESEARCH INSTITUTE
              OS Homo sapiens (human)
              PN JP 2002017375-A/2713
              PD 22-JAN-2002

FEATURES
source
CDS
  1..2359
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAF85958.1"
    /db_xref="GI:45502668"
    /translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSS
    APFLASAVAPQPPDQCPCSEAAATKVCNRLNTEVPTDLPAYVNLFTGNQ
    LAVLPAGAFARRPPLAELALNLSGSLRDEVRAGAFELPSRLDLSHNPLADLSPF
    AFSGSNASVSAPSPPLVELILNHI VPPEDSRQNRSGFMVVAALLAGRALQGLRLLELA
    SNHFLYLPDVLQALPSLRHLDSLNNLSVSLTYVSNLTHLESLELDNALKVLHNG
    TLAEQLGPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPEKMRNVL
    LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLLVLYLNKGIKKWMNIRDAC
    RDMEGYHYRYEINADPRLNLSNSDV"

ORIGIN
Alignment Scores:
Pred. No.:      7.21      Length:      2053
Score:          46.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:        0

US-10-774-176-11 (1-9) x HS5T40A (1-2053)

QY      1 AsnLeuThrGluValProThrAspLeu 9
|||||
Db      325 AATCTGACCGAGGTGCCACGACCTG 351

RESULT 7
BD127282      2359 bp      DNA      linear      PAT 18-SEP-2002
LOCUS
DEFINITION    Primer for synthesizing full-length cDNA and use thereof.
ACCESSION     BD127282
VERSION       BD127282.1 GI:23222227
KEYWORDS      JP 2002017375-A/2713.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE     1 (bases 1 to 2359)
AUTHORS      Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
              Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
              Koga,H.
              Primers for synthesizing full-length cDNA and use thereof
              Patent: JP 2002017375-A/2713 22-JAN-2002;
              HELIX RESEARCH INSTITUTE
              OS Homo sapiens (human)
              PN JP 2002017375-A/2713
              PD 22-JAN-2002

FEATURES
source
CDS
  1..2359
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAF85958.1"
    /db_xref="GI:45502668"
    /translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSS
    APFLASAVAPQPPDQCPCSEAAATKVCNRLNTEVPTDLPAYVNLFTGNQ
    LAVLPAGAFARRPPLAELALNLSGSLRDEVRAGAFELPSRLDLSHNPLADLSPF
    AFSGSNASVSAPSPPLVELILNHI VPPEDSRQNRSGFMVVAALLAGRALQGLRLLELA
    SNHFLYLPDVLQALPSLRHLDSLNNLSVSLTYVSNLTHLESLELDNALKVLHNG
    TLAEQLGPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPEKMRNVL
    LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLLVLYLNKGIKKWMNIRDAC
    RDMEGYHYRYEINADPRLNLSNSDV"

```

```

PF      07-JUL-2000 JP 2000253172
PI      TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI      ISHII,
PI      YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI      TETSUJI OTSUKI,HISASHI KOGA
PC      C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC      C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT      CDS
      Location/Qualifiers
      (424)..(1572).
      source
      1..2359
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      8.24      Length:      2359
Score:          46.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             6          Gaps:        0

US-10-774-176-11 (1-9) x BD127282 (1-2359)

QY      1 AsnLeuThrGluValProThrAspLeu 9
|||||
Db      664 AATCTGACCGAGGTGCCACGACCTG 690

RESULT 8
CQ782724      2359 bp      DNA      linear      PAT 17-MAR-2004
LOCUS
DEFINITION    Sequence 2864 from Patent EP1396543.
ACCESSION     CQ782724
VERSION       CQ782724.1 GI:45502667
KEYWORDS      Homo sapiens (human)
SOURCE        Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE     1
AUTHORS      Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
              Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
              Koga,H.
              Primers for synthesizing full length cDNA clones and their use
              Patent: EP 1396543-A 2864 10-MAR-2004;
              Research Association for Biotechnology (JP)

FEATURES
source
  1..2359
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAF85958.1"
    /db_xref="GI:45502668"
    /translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSS
    APFLASAVAPQPPDQCPCSEAAATKVCNRLNTEVPTDLPAYVNLFTGNQ
    LAVLPAGAFARRPPLAELALNLSGSLRDEVRAGAFELPSRLDLSHNPLADLSPF
    AFSGSNASVSAPSPPLVELILNHI VPPEDSRQNRSGFMVVAALLAGRALQGLRLLELA
    SNHFLYLPDVLQALPSLRHLDSLNNLSVSLTYVSNLTHLESLELDNALKVLHNG
    TLAEQLGPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPEKMRNVL
    LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLLVLYLNKGIKKWMNIRDAC
    RDMEGYHYRYEINADPRLNLSNSDV"

ORIGIN
Alignment Scores:
Pred. No.:      8.24      Length:      2359
Score:          46.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             6          Gaps:        0

US-10-774-176-11 (1-9) x BD127282 (1-2359)

QY      1 AsnLeuThrGluValProThrAspLeu 9
|||||
Db      664 AATCTGACCGAGGTGCCACGACCTG 690

RESULT 8
CQ782724      2359 bp      DNA      linear      PAT 17-MAR-2004
LOCUS
DEFINITION    Sequence 2864 from Patent EP1396543.
ACCESSION     CQ782724
VERSION       CQ782724.1 GI:45502667
KEYWORDS      Homo sapiens (human)
SOURCE        Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homnidae; Homo.
REFERENCE     1
AUTHORS      Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
              Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
              Koga,H.
              Primers for synthesizing full length cDNA clones and their use
              Patent: EP 1396543-A 2864 10-MAR-2004;
              Research Association for Biotechnology (JP)

FEATURES
source
  1..2359
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAF85958.1"
    /db_xref="GI:45502668"
    /translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSS
    APFLASAVAPQPPDQCPCSEAAATKVCNRLNTEVPTDLPAYVNLFTGNQ
    LAVLPAGAFARRPPLAELALNLSGSLRDEVRAGAFELPSRLDLSHNPLADLSPF
    AFSGSNASVSAPSPPLVELILNHI VPPEDSRQNRSGFMVVAALLAGRALQGLRLLELA
    SNHFLYLPDVLQALPSLRHLDSLNNLSVSLTYVSNLTHLESLELDNALKVLHNG
    TLAEQLGPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPEKMRNVL
    LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLLVLYLNKGIKKWMNIRDAC
    RDMEGYHYRYEINADPRLNLSNSDV"

```

```

Pred. No.: 8.24 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x CQ782724 (1-2359)

Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 664 AATCTGACCGAGGTGCCCGACCTG 690

RESULT 9
AK074786 2359 bp mRNA linear PRI 03-SEP-2002
LOCUS Homo sapiens cDNA FLJ30305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AK074786
VERSION AK074786.1 GI:22760460
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2359)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1. .2359
Location/Qualifiers
/organism="Homo sapiens"
/db_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/notes="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"
ORIGIN
Alignment Scores:
Pred. No.: 8.24 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x AK074786 (1-2359)
Qy 1 AsnLeuThrGluValProThrAspLeu 9

Pred. No.: 8.24 Length: 2359
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x CQ782726 (1-2361)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 666 AATCTGACCGAGGTGCCCGACCTG 692

RESULT 11
CQ782726 2361 bp DNA linear PAT 17-MAR-2004
LOCUS Sequence 2866 from Patent EP1396543.
ACCESSION CQ782726
VERSION CQ782726.1 GI:45502669
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,

```

```

Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
Primers for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 2866 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES             source
    source
    1. .2361
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
    426..1688
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAF85961.1"
        /db_xref="GI:45502670"
        /translation="MPGCGSRGPAAGDGRLERLARLALVLLGWSSSPPTSSASSFSSS
        AFLPAVAQAQPLPQPCALCESEARTVKCNRLNTEVPTDLPAYVNRNLTGNQ
        LAVLPAGAPRPPLAEALALNLSGSRLEVRAGAFELPSPQLDLSHNPDLSPF
        AFGSGNASVAPSLVELILNHTVPPEDEQRNRSFEGMVVAALLAGRALQGLRLLELA
        SNHFLYLPDVLQALPSLHLDLSNNLSVSLTVSPFNLTLSLESLHEDNALKVLHNG
        TLAEGLGPHIRVFLPNNPWCDCMADVMYTLKETEVCGKDRLTCAYPEKMRNRL
        LELNSADLDCDPIPLPSLQTSYVPLGIVLALIGALFLLVLYLNKGIKKWMHNRDAR
        RDHMEGYHYRYEINADPRLTNLSNSDV"
    CDS
    2361
        /codon_start=1
        /protein_id="CAF85961.1"
        /db_xref="GI:45502670"
        /translation="MPGCGSRGPAAGDGRLERLARLALVLLGWSSSPPTSSASSFSSS
        AFLPAVAQAQPLPQPCALCESEARTVKCNRLNTEVPTDLPAYVNRNLTGNQ
        LAVLPAGAPRPPLAEALALNLSGSRLEVRAGAFELPSPQLDLSHNPDLSPF
        AFGSGNASVAPSLVELILNHTVPPEDEQRNRSFEGMVVAALLAGRALQGLRLLELA
        SNHFLYLPDVLQALPSLHLDLSNNLSVSLTVSPFNLTLSLESLHEDNALKVLHNG
        TLAEGLGPHIRVFLPNNPWCDCMADVMYTLKETEVCGKDRLTCAYPEKMRNRL
        LELNSADLDCDPIPLPSLQTSYVPLGIVLALIGALFLLVLYLNKGIKKWMHNRDAR
        RDHMEGYHYRYEINADPRLTNLSNSDV"
    ORIGIN
    Alignment Scores:
    Pred. No.:      8.25      Length:      2361
    Score:          46.00     Matches:      9
    Percent Similarity: 100.0% Conservatives: 0
    Best Local Similarity: 100.0% Mismatches: 0
    Query Match:    100.0% Indels:      0
    DB:             6        Gaps:        0

US-10-774-176-11 (1-9) x CQ782726 (1-2361)

Qy      1 AsnLeuThrGluValProThrAspLeu 9
      |||||
Db      666 AATCTGACCGAGGTGCCCGACCTG 692

RESULT 12
AX961916      2361 bp      DNA      linear      PAT 14-JAN-2004
LOCUS
DEFINITION
Sequence 127 from Patent WO03104277.
ACCESSION
AX961916.1 GI:40881326
KEYWORDS
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1
Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
Stat6 activation gene
Patent: WO 03104277-A 127 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES             source
    source
    1. .2361
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
    426..1688
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAF06467.1"
        /db_xref="GI:40881327"
        /translation="MPGCGSRGPAAGDGRLERLARLALVLLGWSSSPPTSSASSFSSS
        AFLPAVAQAQPLPQPCALCESEARTVKCNRLNTEVPTDLPAYVNRNLTGNQ
        LAVLPAGAPRPPLAEALALNLSGSRLEVRAGAFELPSPQLDLSHNPDLSPF
        AFGSGNASVAPSLVELILNHTVPPEDEQRNRSFEGMVVAALLAGRALQGLRLLELA
        SNHFLYLPDVLQALPSLHLDLSNNLSVSLTVSPFNLTLSLESLHEDNALKVLHNG
        TLAEGLGPHIRVFLPNNPWCDCMADVMYTLKETEVCGKDRLTCAYPEKMRNRL
        LELNSADLDCDPIPLPSLQTSYVPLGIVLALIGALFLLVLYLNKGIKKWMHNRDAR
        RDHMEGYHYRYEINADPRLTNLSNSDV"
    CDS
    2361
        /codon_start=1
        /protein_id="CAF06467.1"
        /db_xref="GI:40881327"
        /translation="MPGCGSRGPAAGDGRLERLARLALVLLGWSSSPPTSSASSFSSS
        AFLPAVAQAQPLPQPCALCESEARTVKCNRLNTEVPTDLPAYVNRNLTGNQ
        LAVLPAGAPRPPLAEALALNLSGSRLEVRAGAFELPSPQLDLSHNPDLSPF
        AFGSGNASVAPSLVELILNHTVPPEDEQRNRSFEGMVVAALLAGRALQGLRLLELA
        SNHFLYLPDVLQALPSLHLDLSNNLSVSLTVSPFNLTLSLESLHEDNALKVLHNG
        TLAEGLGPHIRVFLPNNPWCDCMADVMYTLKETEVCGKDRLTCAYPEKMRNRL
        LELNSADLDCDPIPLPSLQTSYVPLGIVLALIGALFLLVLYLNKGIKKWMHNRDAR
        RDHMEGYHYRYEINADPRLTNLSNSDV"
    ORIGIN
    Alignment Scores:
    Pred. No.:      8.25      Length:      2361
    Score:          46.00     Matches:      9
    Percent Similarity: 100.0% Conservatives: 0
    Best Local Similarity: 100.0% Mismatches: 0
    Query Match:    100.0% Indels:      0
    DB:             6        Gaps:        0

US-10-774-176-11 (1-9) x CQ782726 (1-2361)

Qy      1 AsnLeuThrGluValProThrAspLeu 9
      |||||
Db      666 AATCTGACCGAGGTGCCCGACCTG 692

RESULT 13
AX074790      2361 bp      mRNA      linear      PRI 09-JUL-2005
LOCUS
DEFINITION
Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION
AX074790
VERSION
AK074790.1 GI:22760466
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1
Otsuki,T., Ota,T., Nishikawa,T., Hayashi,K., Suzuki,Y.,
Yamamoto,J., Wakamatsu,A., Kimura,K., Sakamoto,K., Hatano,N.,
Kawai,Y., Ishii,S., Saito,K., Kojima,S., Sugiyama,T., Ono,T.,
Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A.,
Okumura,K., Nagai,K., Sugano,S. and Isogai,T.
Signal Sequence and Keyword Trap in silico for Selection of
Full-Length Human cDNAs Encoding Secretion or Membrane Proteins
from Oligo-Capped cDNA Libraries
DNA Res. 12, 117-126 (2005)
JOURNAL
2
Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuhio,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2361)
Isogai,T. and Otsuki,T.
Direct Submission
JOURNAL
TITLE
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES             source
    source
    1. .2361
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="NT2RP2000903"
        /cell_line="NT2"
        /cell_type="teratocarcinoma"
        /clone_lib="NT2RP2"
        /note="cloning vector: pME18SFL3
        mRNA from NT2 neuronal precursor cells after 2-weeks
  
```

RDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:

Pred. No.: 8.25 Length: 2361
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x AX961916 (1-2361)

Qy 1 AsnLeuThrGluValProThrAspLeu 9

|||||
 Db 666 AATCTGACCGAGGTGCCCGACCTG 692

RESULT 13

AX074790

LOCUS

DEFINITION

Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar

to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.

ACCESSION

AX074790

VERSION

AK074790.1 GI:22760466

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

REFERENCE

1

Otsuki,T., Ota,T., Nishikawa,T., Hayashi,K., Suzuki,Y.,

Yamamoto,J., Wakamatsu,A., Kimura,K., Sakamoto,K., Hatano,N.,

Kawai,Y., Ishii,S., Saito,K., Kojima,S., Sugiyama,T., Ono,T.,

Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A.,

Okumura,K., Nagai,K., Sugano,S. and Isogai,T.

Signal Sequence and Keyword Trap in silico for Selection of

Full-Length Human cDNAs Encoding Secretion or Membrane Proteins

from Oligo-Capped cDNA Libraries

DNA Res. 12, 117-126 (2005)

JOURNAL

2

Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,

Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,

Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,

Kojima,S., Nagahari,K., Masuhio,Y., Ono,T., Okano,K., Yoshikawa,Y.,

Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and

Ninomiya,K.

NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 2361)

Isogai,T. and Otsuki,T.

Direct Submission

JOURNAL

TITLE

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology; cDNA library construction:

Institute of Medical Science, University of Tokyo, Laboratory of

Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass

sequencing and clone selection: Helix Research Institute (supported

by Japan Key Technology Center etc.).

FEATURES

source

1. .2361

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP2000903"

/cell_line="NT2"

/cell_type="teratocarcinoma"

/clone_lib="NT2RP2"

/note="cloning vector: pME18SFL3

mRNA from NT2 neuronal precursor cells after 2-weeks

```

retinoic acid (RA) induction"

ORIGIN
Alignment Scores:
Pred. No.:      8.25      Length:      2361
Score:          46.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:      0

US-10-774-176-11 (1-9) x AK074790 (1-2361)

QY      1 AsnLeuThrGluValProThrAspLeu 9
Db      666 AATCTGACCGAGGTGCCACGACCTG 692

RESULT 14
BC037161
LOCUS
DEFINITION
Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317
IMAGE:4138906), complete cds.
ACCESSION
BC037161
VERSION
BC037161.2 GI:33872201
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1 (bases 1 to 2379)
Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvedchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
JOURNAL
PUBMED
2 (bases 1 to 2379)
Strausberg, R.
Direct Submission
Submitted (03-SEP-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:22713382.
Contact: MGC help desk
Email: ggapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing Arrayed by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov
Series: IRAL Plate: 26 Row: m Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5729717.
FEATURES
source
Location/Qualifiers
1..2379
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:15317 IMAGE:4138906"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2379
/gene="TPBG"
/note="synonyms: M6P1, 5T4-AG, 5T4"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"
427..1689
/gene="TPBG"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="AAH37161.1"
/db_xref="GI:22713383"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"
/translation="MPGCSRGPAAGDGLRLRLALVLVHVVSSSPSSPSSSS
APFLASVAQPIPLDQPCALCESEAAATVKVNRNTEVPTDLPAYVYRNLFTGNQ
LAVLPAGAFARPPPLAALNLSGRSLDEVRAGAFELPSLRQLDLSHNPADLSPPF
AFSGNSASVAPSPLVELILNHIVPEDQRNRSFEGMVVAALLAGLALQGLRLLELA
SNFELYLPDRVLAQLPSLRHLDLNNLSVLSVTYSFRNLTLESLEHLEDAKVLHNG
TLASLQGLPHLRVFLDNNPNWCDCHMDMVTWKEVTVQKDRLTCAYPEKMRNRL
LELNSADDLCPILPPSPQTSYVIGIVLALIGALFLVLYLNRKIKKMMNIRDAC
RDHMEGYHYRYEINADRLTLNLSNSDV"
ORIGIN
Alignment Scores:
Pred. No.:      8.31      Length:      2379
Score:          46.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:      0

US-10-774-176-11 (1-9) x BC037161 (1-2379)

QY      1 AsnLeuThrGluValProThrAspLeu 9
Db      667 AATCTGACCGAGGTGCCACGACCTG 693

RESULT 15
AB168308
LOCUS
DEFINITION
Macaca fascicularis testis cDNA clone: Qtsa-11109, similar to human
trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.
ACCESSION
AB168308
VERSION
AB168308.1 GI:67967899
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Macaca fascicularis (crab-eating macaque)
ORGANISM
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

```

1 International consortium for macaque cDNA sequencing and analysis.
DNA sequences of macaque genes expressed in brain or testis and its
evolutionary implications
Unpublished

2 Osada, N., Hirata, M., Tanuma, R., Kusuda, J., Hida, M., Suzuki, Y.,
Sugano, S., Gojobori, T., Shen, J.-C.-K., Wu, C.-I. and Hashimoto, K.
Substitution rate and structural divergence of 5'UTR evolution:
Comparative analysis between human and cynomolgus monkey cDNAs
Unpublished
3 (bases 1 to 2714)

Hashimoto, K., Kusuda, J. and Sugano, S.

Direct Submission
Submitted (18-MAR-2004) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan
(E-mail: khashim@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
The International Consortium for macaque cDNA sequencing and
analysis consists of: Department of Virology and Human Genome
Center, Institute of Medical Science, The University of Tokyo,
Tokyo, Japan; Division of Genetic Resources, National Institute of
Infectious Diseases of Japan, Tokyo, Japan; National Health
Research Institute, Taipei, Taiwan; Institute of Molecular Biology,
Academia Sinica, Taipei, Taiwan; Department of Ecology & Evolution,
University of Chicago, Chicago, IL, USA; Center for Information
Biology, National Institute of Genetics of Japan, Mishima, Japan.
Clone distribution: clone distribution information can be found at:
http://www.nih.go.jp/yoken/genebank/
Lab host: TOP10

Vector: pME18S-FL3 (Acc. No. AB009864)

R. Site1: DraIII (CACATGTCG)

R. Site2: DraIII (CACATGTCG)

Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGCGCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5 kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by oligo-capping method. Libraries were made from:

OqcB: cerebellum cortex

OmpA: parietal lobe

QtrA: temporal lobe right

QfIA: frontal lobe left

QmoA: medulla oblongata

QbsA: brain stem

QorA: occipital lobe right

QtsA: testis

Custom primers were used for 5' and 3'-end sequencing. The
full-insert sequencing was done by primer-walking method using ABI
DNA sequencer.

FEATURES
source

Location/Qualifiers
1..2714
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="QtsA-11109"
/sex="male"
/clone_lib="macaque cDNA library QtsA"
/dev_stage="adult"
764..2026

CDS

/notes="unnamed protein product; Homo sapiens trophoblast
glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3"
/codon_start=1
/protein_id="BA200432.1"
/db_xref="GI:67967900"
/translation="MFGCGSRGPAAGDGRRLRLRLALVLLGWSSSSSTSSASSSSSS
APFLAASAAQPPPLDQCPCALCECSAAATVKCNRLNTEVPTDLPVLRNLFITNGQ
LAVLPAGAPRFLPALANLSGRLDVRGGAFFHLPSLRDLSDHSLAYLSFP
AFSGSNASTSAPSPVELILNHPDPRKQNSFEQVAAALVAGRALQLHLELA
SNHFLYLPRLVLAQLPSRLYLDLSNNSLVSLTVYSPFNLTLSLSLHLEDNALVKYLHG

ORIGIN

Alignment Scores:
Pred. No.: 9.42 Length: 2714
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x AB168308 (1-2714)

Qy 1 AsnLeuThrGluValProThrAspLeu 9

Db 1004 ANCTGACCGAGGTGCCACGGACCTG 1030

RESULT 16

HSA012159

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

misc_binding

misc_binding

gene

exon

intron

exon

CDS

Location/Qualifiers

1..5551

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

2698..2703

/bound_moiety="Spl"

2704..2709

/bound_moiety="Spl"

2716..5400

/gene="5T4"

/evidence="experimental"

2801..3092

/gene="5T4"

/evidence="experimental"

3093..5400

/gene="5T4"

/evidence="experimental"

3431..4693

/gene="5T4"

/codon_start=1

/product="5T4 oncofetal trophoblast glycoprotein"

/protein_id="CAA09930.1"

/db_xref="GI:3805947"

/db_xref="GOA:Q13641"

/db_xref="InterPro:IPR000372"

/db_xref="InterPro:IPR000483"

TLAEQLPHVRVFLDNNPMVCDCHMADMTLTKOTGVQKORLTCAFPKMRNRL
LENSADLDCCDPLPSPISQTSYVFLGLVLAIGLFLVLVLRKGIKKWMHNRDAD
RDMEGYHYRYEINADPRLTNLSSDV"

```

/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCSRGPAAGDGLRLRLARLALVLLGWSSSPTSSASSFSS
APFLASVSAQPPPLDQCPCALCESEARTVKCVNRNLTEVPTDLPAYVNRNLTGNQ
LAVLPAGAFARRPPLAEALNLISGRLDEVRAGAFHLPSLRQLDLSHNPLADLSPF
AFSGSNASVSPPLVELILNHIYVPEDEQRNSFEGMVVAALLAGRALQGLRLELA
SNHFLYLPDRVLAQLPSRLHDLNNSLSVLTYSFRNLTHLESLEDNALKVLHNG
TLAELQGLPHIRVFLDNNPWCDHMADMTWLKETEVQCKDRLTCAYPEKMRNRL
LELNSADLDCDPLPSPSLQTSYVFLGIVLALIGNAFLVLLVLYNRKGIKWMHNRDAC
RDMEGYHYRYEINADPRLTNLSNSDV"
3431..3516
/gene="574"
/mat_peptide
3517..4690
/gene="574"
/product="574 oncofetal trophoblast glycoprotein"
5331..5336
/gene="574"
/polyA_signal
5380..5385
/gene="574"

```

sig_peptide

mat_peptide

polyA_signal

polyA_signal

ORIGIN

```

Alignment Scores:
Pred. No.:      18.7      Length:      5551
Score:          46.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      100.0%      Indels:      0
DB:               8        Gaps:      0

```

US-10-774-176-11 (1-9) x HSA012159 (1-5551)

QY 1 AnLeuThGluValProThrAspLeu 9

Db 3671 AATCTGACCGAGTGCACCGACCTG 3697

RESULT 17

HSJ492P14

LOCUS

DEFINITION

Human DNA sequence from clone RP3-492P14 on chromosome 6q13-15
 Contains a single stranded DNA binding protein pseudogene, the TPBG
 gene for trophoblast glycoprotein (574-AG) and a CpG island,
 complete sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Al121977
 AL121977.11 GI:11863678
 HTG; CpG island; TPBG.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 121909)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
 Clone requests: Clonerequest@sanger.ac.uk
 On Dec 15, 2000 this sequence version replaced gi:11558491.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6
 RP3-492P14 is from the library RPc1-3 constructed by the group of
 Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2

----- Genome Center
 Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one subclone, and the assembly was confirmed by restriction digest,
 except on the rare occasion of the clone being a YAC.

FEATURES

source

Location/Qualifiers

1..121909

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="6"

/map="q13-15"

/clone="RP3-492P14"

/clone_lib="RPC1-3"

misc_feature

100

/note="Clone right end: RP1-93K22"

/complement(10004..10982)

/locus_tag="RP3-492P14.2-001"

/pseudo

/complement(10004..10982)

/locus_tag="RP3-492P14.2-001"

/note="match: proteins: P81877 Q99LX9 Q9BWM6 Q9CYZ8 Q9B6L4
Q9P038 Q9Y4T7"

/pseudo

/codon_start=1

misc_feature

86539

/note="Clone left end: RP1-90G1"

/complement(109639..116836)

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

/join(109639..109916,110631..116836)

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

/product="trophoblast glycoprotein"

/note="match: ESTs: AA149121 AA152323 AA565852 AA643734
AL544610 AW471072 AW662538 BE260089 BF306457 BF306926
BF314984 BI196133 B1562387 BM089633 BM670613

match: cDNAs: AJ420536.1 Z29083.1"

110970..112232

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

/standard_name="OTTHUMP00000016786"

/note="match: proteins: Q13641 Q9QYD9 Q9ZOL0"

/codon_start=1

/product="trophoblast glycoprotein"

/protein_id="CAI21546.1"

/db_xref="GI:56203539"

/db_xref="Genbank:12004"

/db_xref="GOA:Q13641"

/db_xref="InterPro:IPR000372"

/db_xref="InterPro:IPR000483"

/db_xref="InterPro:IPR001611"

/db_xref="InterPro:IPR003591"

/db_xref="UniProt/TREMBL:Q13641"

/translation="MPGCSRGPAAGDGLRLRLARLALVLLGWSSSPTSSASSFSS
 APFLASVSAQPPPLDQCPCALCESEARTVKCVNRNLTEVPTDLPAYVNRNLTGNQ
 LAVLPAGAFARRPPLAEALNLISGRLDEVRAGAFHLPSLRQLDLSHNPLADLSPF
 AFSGSNASVSPPLVELILNHIYVPEDEQRNSFEGMVVAALLAGRALQGLRLELA
 SNHFLYLPDRVLAQLPSRLHDLNNSLSVLTYSFRNLTHLESLEDNALKVLHNG
 TLAELQGLPHIRVFLDNNPWCDHMADMTWLKETEVQCKDRLTCAYPEKMRNRL
 LELNSADLDCDPLPSPSLQTSYVFLGIVLALIGNAFLVLLVLYNRKGIKWMHNRDAC
 RDMEGYHYRYEINADPRLTNLSNSDV"

116817..116822

polyA_signal

/gene="TPBG"

/locus_tag="RP3-492P14.1-001"

116836

polyA_site

/gene="TPBG"

```
/locus_tag="RP3-492P14.1-001"
misc_feature 121909
/notes="Clone_right_end: RP3-492P14"

ORIGIN
Alignment Scores:
Pred. No.: 358 Length: 121909
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x H5J492P14 (1-121909)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 111210 AATCTGACCGAGGTGCCCGGACCTG 111236

RESULT 18
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS Felis sp.
SOURCE Felis sp.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE 1
AUTHORS Myer, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. .1260
/mol_type="unassigned DNA"
/db_xref="taxon:9687"

ORIGIN
Alignment Scores:
Pred. No.: 56.9 Length: 1260
Score: 41.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.1% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x AX467373 (1-1260)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 238 AACCTGACCGAGGTGCCCGGACCTG 264

RESULT 19
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS Felis catus (cat)
SOURCE Felis catus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE 1
AUTHORS Carroll, M.M., Kingsman, S.M. and Redchenko, I.M.
TITLE MHC class I peptide epitopes from the human 5t4 tumor-associated
antigen

JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. .1260
/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN
Alignment Scores:
Pred. No.: 56.9 Length: 1260
Score: 41.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.1% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x AX821548 (1-1260)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 238 AACCTGACCGAGGTGCCCGGACCTG 264

RESULT 21
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kingsman, A.O., Kingsman, S.M., Bebbington, C.R., Carroll, M.W.,
```

TITLE Ellard, F.M. and Myers, K.A.
 JOURNAL Antibodies
 Patent: WO 0136486-A 14 25-MAY-2001;
 Oxford Biomedica (UK) Limited (GB)
 FEATURES Location/Qualifiers
 source 1..1263
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /db_xref="514"

ORIGIN

Alignment Scores: 1263
 Pred. No.: 57 Length: 8
 Score: 41.00 Matches: 0
 Percent Similarity: 88.9% Conservat: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 89.1% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x AX149553 (1-1263)

QY 1 AsnLeuThrGluValProThrAspLeu 9
 |||||
 Db 241 AACCTGACCGAGGTGCCCGGACCTG 267

RESULT 22

LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
 DEFINITION Sequence 1 from Patent WO0238612.
 ACCESSION AX467371
 VERSION AX467371.1 GI:21900602
 KEYWORDS
 SOURCE Canis sp.
 ORGANISM Canis sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.

REFERENCE 1
 AUTHORS Myers, K., Drury, N. and Carroll, M.
 TITLE Polypeptide
 JOURNAL Patent: WO 0238612-A 1 15-MAY-2002;
 Oxford Biomedica (UK) Limited (GB)
 FEATURES Location/Qualifiers
 source 1..1263
 /organism="Canis sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:9616"

ORIGIN

Alignment Scores: 1263
 Pred. No.: 57 Length: 8
 Score: 41.00 Matches: 0
 Percent Similarity: 88.9% Conservat: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 89.1% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x AX467371 (1-1263)

QY 1 AsnLeuThrGluValProThrAspLeu 9
 |||||
 Db 241 AACCTGACCGAGGTGCCCGGACCTG 267

RESULT 23

AP006878_12/c
 WPCOMMENT
 Sequence split into 21 fragments LOCUS AP006878 Accession AP006878

Fragment Name	Begin	End
AP006878_00	1	110000
AP006878_01	100001	210000
AP006878_02	200001	310000
AP006878_03	300001	410000

AP006878_04 400001 510000
 AP006878_05 500001 610000
 AP006878_06 600001 710000
 AP006878_07 700001 810000
 AP006878_08 800001 910000
 AP006878_09 900001 1010000
 AP006878_10 1000001 1110000
 AP006878_11 1100001 1210000
 AP006878_12 1200001 1310000
 AP006878_13 1300001 1410000
 AP006878_14 1400001 1510000
 AP006878_15 1500001 1610000
 AP006878_16 1600001 1710000
 AP006878_17 1700001 1810000
 AP006878_18 1800001 1910000
 AP006878_19 1900001 2010000
 AP006878_20 2000001 2088737
 Continuation (13 of 21) of AP006878 from base 1200001 (AP006878 Thermococcus kodakarensis)

Alignment Scores:

Pred. No.: 4.08e+03 Length: 110000
 Score: 41.00 Matches: 7
 Percent Similarity: 100.0% Conservat: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 89.1% Indels: 0
 DB: 1 Gaps: 0

US-10-774-176-11 (1-9) x AP006878_12 (1-110000)

QY 1 AsnLeuThrGluValProThrAspLeu 9
 |||||
 Db 58502 AATATCACCGAGGTGCCCACTGACCTC 58476

RESULT 24

LOCUS AC137462 187783 bp DNA linear HTG 20-NOV-2002
 DEFINITION Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
 ***, 6 unordered pieces.

ACCESSION

AC137462.1 GI:25138601

HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE

1 (bases 1 to 187783)
 Muzny, D., Marie, Metzker, M., Lee, A., Adams, S., Alder, J.,
 Allen, C., Allen, H., Alebrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biawalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhewa, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaoketeme, O., Okwuonu, G., Olarunpusagbon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoeh, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quintero, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shadman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J.,
 Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G. and Gibbs, R.A.
 Direct Submission
 Unpublished

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

2 (bases 1 to 187783)
 Rat Genome Sequencing Consortium.
 Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: KZ1Q

Center clone name: CH230-unknown

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 149899 bases at least Q40

Consensus quality: 155975 bases at least Q30

Consensus quality: 160292 bases at least Q20

Estimated insert size: 153789; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N. But the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 40744: contig of 40744 bp in length
 * 40745 40844: gap of unknown length
 * 40845 180556: contig of 139712 bp in length
 * 180557 180656: gap of unknown length

* 180657 182042: contig of 1386 bp in length
 * 182043 182142: gap of unknown length
 * 182143 183473: contig of 1331 bp in length
 * 183474 183573: gap of unknown length
 * 183574 183510: contig of 1737 bp in length
 * 183511 185410: gap of unknown length
 * 185411 187783: contig of 2373 bp in length.

FEATURES

source
 1. 187783
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-unknown"
 36541..37777
 /note="wgs contig"
 40745..40844
 /estimated_length=unknown
 76822..80300
 /note="wgs contig"
 107643..109294
 /note="wgs contig"
 147795..149099
 /note="wgs contig"
 169577..171056
 /note="wgs contig"
 180557..180656
 /estimated_length=unknown
 182043..182142
 /estimated_length=unknown
 183474..183573
 /estimated_length=unknown
 185311..185410
 /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 6.81e+03 Length: 187783
 Score: 41.00 Matches: 8
 Percent Similarity: 88.9% Conservative: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 89.1% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-11 (1-9) x AC137462 (1-187783)

Qy 1 AsnLeuThrGluValProThrAspIeu 9

Db 52513 AACCTAAGTAGGGTACCTACTGACTTG 52539

RESULT 25

AC158075/c

LOCUS AC158075 225405 bp DNA linear HTG 01-JUL-2005
 DEFINITION Bos taurus clone CH240-53E16, *** SEQUENCING IN PROGRESS ***, 26
 unordered pieces.

AC158075

AC158075.2 GI:68267879

HTG; HTGS_PHASRI; HTGS_DRAFT; HTGS_ENRICHED.

Bos taurus (cow)

ORGANISM

Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 225405)

Muzny, D., Marie, Metzker, M., Lee, A., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alabrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayvagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biwalto, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gearegroat, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, N., Hognes, M., Hollins, B., Howells, S., Huly, S., Hume, J., Idiebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensueva, J., Loulsged, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahndartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, J., Savary, G., Scherer, S., Scott, G., Shateman, S., Shen, H., Shetty, W., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinfeld, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, I., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 225405)
Worley, K.C.
Direct Submission
Submitted (04-MAR-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 225405)
Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 28, 2005 this sequence version replaced gi:60498847.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: FDBK

Center clone name: CH240-53E16
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 212622 bases at least Q40
Consensus quality: 214515 bases at least Q30
Consensus quality: 216171 bases at least Q20
Estimated insert size: 216549; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 4233: contig of 4233 bp in length
4234 5569: gap of 1336 bp
5570 7744: contig of 2175 bp in length
7745 8270: gap of 526 bp
8271 10776: contig of 2506 bp in length
10777 11091: gap of 315 bp
11092 13887: contig of 2796 bp in length
13888 13937: gap of 50 bp
13938 25770: contig of 11833 bp in length
25771 39559: contig of 13689 bp in length
39560 39609: gap of 50 bp
39610 43339: contig of 3730 bp in length
43340 44663: gap of 1324 bp
44664 49227: contig of 4564 bp in length
49228 49278: gap of 50 bp
49279 80346: contig of 31069 bp in length
80347 80396: gap of 50 bp
80397 105366: contig of 24970 bp in length
105367 106918: gap of 1552 bp
106919 125551: contig of 18633 bp in length
125552 125601: gap of 50 bp
125602 128929: contig of 3328 bp in length
128930 129029: gap of unknown length
129030 130822: contig of 1792 bp in length
130823 130911: gap of 90 bp
130912 139228: contig of 8317 bp in length
139229 139278: gap of 50 bp
139279 143067: contig of 3789 bp in length
143068 143117: gap of 50 bp
143118 161612: contig of 18495 bp in length
161613 161662: gap of 50 bp
161663 176870: contig of 15208 bp in length
176871 176920: gap of 50 bp
176921 187778: contig of 10858 bp in length
187779 187828: gap of 50 bp
187829 214857: contig of 27029 bp in length
214858 214957: gap of unknown length
214958 216517: contig of 1560 bp in length
216518 216617: gap of unknown length
216618 218043: contig of 1426 bp in length
218044 218143: gap of unknown length
218144 219381: contig of 1238 bp in length
219382 219481: gap of unknown length
219482 221188: contig of 1707 bp in length
221189 221288: gap of unknown length
221289 222411: contig of 1123 bp in length
222412 222511: gap of unknown length
222512 223707: contig of 1196 bp in length
223708 223807: gap of unknown length
223808 225405: contig of 1598 bp in length.
Location/Qualifiers
1. 225405
/organism="Bos taurus"

FEATURES
source

```

/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-53816"
4234..5569
/estimated_length=1336
7745..8270
/estimated_length=526
10777..11091
/estimated_length=315
13888..13937
/estimated_length=50
25771..25870
/estimated_length=unknown
39560..39609
/estimated_length=50
43340..44663
/estimated_length=1324
49228..49277
/estimated_length=50
80347..80396
/estimated_length=50
105367..106918
/estimated_length=1552
125552..125601
/estimated_length=50
128930..129029
/estimated_length=unknown
130822..130911
/estimated_length=90
139229..139278
/estimated_length=50
143068..143117

```

```

Alignment Scores:
Pred. No.: 8.11e-03 Length: 225405
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.1% Indels: 0
DB: 14 Gaps: 0

```

US-10-774-176-11 (1-9) x AC158075 (1-225405)

```

QY 1 AenLeuThrcLuvAlProThraPLeu 9
DB 54769 AACTTGACTAAGATTCCACAGATTTA 54743

```

```

RESULT 26
AC115158/c
LOCUS AC115158 241280 bp DNA linear HTG 19-NOV-2002
DEFINITION Rattus norvegicus clone CH230-171M15, *** SEQUENCING IN PROGRESS
ACCESSION AC115158
VERSION AC115158.5 GI:25072995
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 241280)
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooke,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,G.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,X., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davilla,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Drapar,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

```

```

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhea,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,N., Murphy,M., Naif,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwokedeme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,P., Poinexter,A., Popovic,D., Primus,B., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 241280)
Worley,K.C.
Direct Submission
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 241280)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23321705.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOZQ
Center clone name: CH230-171M15
----- Summary Statistics

```

Assembly program: Phrap; version 0.990329
 Consensus quality: 148835 bases at least Q40
 Consensus quality: 153713 bases at least Q30
 Consensus quality: 157347 bases at least Q20
 Estimated insert size: 147852; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 197836: contig of 197836 bp in length
 * 197837 197936: gap of unknown length
 * 197937 232527: contig of 34591 bp in length
 * 232528 232627: gap of unknown length
 * 232628 232667: contig of 1040 bp in length
 * 232668 233767: gap of unknown length
 * 233768 234789: contig of 1022 bp in length
 * 234790 234889: gap of unknown length
 * 234890 236876: contig of 1987 bp in length
 * 236877 236976: gap of unknown length
 * 236977 238939: contig of 1963 bp in length
 * 238940 239039: gap of unknown length
 * 239040 241280: contig of 2241 bp in length.

FEATURES

source

Location/Qualifiers
 1. 241280
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-171M15"
 28810..30060
 /notes="wgs contig"
 30111..31608
 /notes="wgs contig"
 58282..59750
 /notes="wgs contig"
 84302..85344
 /notes="wgs contig"
 116977..118507
 /notes="wgs contig"
 159135..150674
 /notes="wgs contig"

gap
 197837..197936
 /estimated_length=unknown
 gap
 232528..232627
 /estimated_length=unknown
 gap
 233668..233767
 /estimated_length=unknown
 gap
 234790..234889
 /estimated_length=unknown
 gap
 236877..236976
 /estimated_length=unknown
 gap
 238940..239039
 /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 8 65e+03 Length: 241280
 Score: 41.00 Matches: 8
 Percent Similarity: 88.9% Conservatives: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 89.1% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-11 (1-9) x AC115158 (1-241280)

QY 1 AsnLeuThrGluValProThrAspLeu 9

Db 217408 AACCTACTAGGTTACTACTGTTG 217382
 RESULT 27
 AC095339/c
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-111N19, *** SEQUENCING IN PROGRESS
 AC095339 267375 bp DNA linear HTG 10-MAY-2003
 *** 17 unordered pieces.
 AC095339
 VERSION
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Rattus.
 1 (bases 1 to 267375)
 REFERENCE
 1 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alebrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gregeorge, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhewa, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
 Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plummer, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. D.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
 Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlecyk, R., Woodden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstein, G., and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 267375)
 REFERENCE
 1 Worley, K. C.
 Direct Submission
 Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One


```

ACCESSION BV561288
VERSION BV561288.1 GI:62452309
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Homiidae; Pan.
REFERENCE 1 (bases 1 to 701)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
        Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the
JOURNAL Human Genome
COMMENT Unpublished (2005)

Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 701
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to
the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps
(Donald,Karlén,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the
western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in
chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred
score >= 20, at least 30% of its base calls must satisfy
SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in
the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25)
bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not
consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied
to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between
two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To
exclude alignment between two
copies of a single read, comparisons between two reads that share
95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were
placed at the same locus
of human genome) were discarded.

FEATURES
    source
        1..701
        /organism="Pan troglodytes verus"
        /mol_type="genomic DNA"
        /sub_species="verus"
        /db_xref="taxon:37012"
        /clone_l1b="Clint"
        <1..>701

ORIGIN
Alignment Scores:
Pred. No.: 53.8 Length: 701
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

```

```

Query Match: 87.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-11 (1-9) x BV561288 (1-701)

QY 2 LeuThrGluValProThrAspLeu 9
    |||||
    658 CTTACGGAGGTGCTACGGACCTC 681

RESULT 29
BV561288
LOCUS
DEFINITION
    qxg68d12.g1 Clint Pan troglodytes verus STS genomic, sequence
    tagged site.
ACCESSION BV561288.1 GI:62455150
VERSION BV561288.1
KEYWORDS STS.
SOURCE Pan troglodytes verus
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Homiidae; Pan.
REFERENCE 1 (bases 1 to 804)
AUTHORS Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and
        Jaffe,D.B.
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the
JOURNAL Human Genome
COMMENT Unpublished (2005)

Contact: Michael C. Zody
Broad Institute of MIT and Harvard
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172580933
Fax: 6172580903
Email: mczody@broad.mit.edu
Primer A: No sequence submitted
Primer B: No sequence submitted
STS size: 804
Protocol:
23,021,928 chimpanzee whole genome shotgun reads were aligned to
the Human genome NCBI
Build 34 (hg16,July 2003). Chimp WGS reads were from 9 donors,
including Clint (Pan
troglodytes verus), 3 other Pan troglodytes verus chimps
(Donald,Karlén,Yvonne), 3 Pan
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps
of unknown origin
(Gon,Unknown Chimp). Common names: Pan troglodytes verus is the
western chimp and Pan
troglodytes troglodytes is the central chimp. To be included in
chimpanzee SNP discovery, a
read must be at least 500bp in length, at least 50% of its base
calls must have Phred
score >= 20, at least 30% of its base calls must satisfy
SNQS(30,25)(single strand NQS, the
base in question has Phred score >= 30, the surrounding 10 bases in
the read have Phred
score >= 25), and the read must have at least 200 bp SNQS(30,25)
bases. Reads not uniquely
placed in the genome and read pairs whose two ends were not
consistently placed were
discarded. After above filtering, NQS(30,25) standard was applied
to all pairs of
overlapping reads to call NQS bases and SNPs. Alignments (between
two reads) with less
than 100 NQS bases or with SNP rate > 0.01 were discarded. To
exclude alignment between two
copies of a single read, comparisons between two reads that share
95% of their genome
alignments (>=95% bases of read A and >=95% bases of read B were
placed at the same locus
of human genome) were discarded.

FEATURES
    Location/Qualifiers

```

```

source      1. .804
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1. .>804

```

```

STS
ORIGIN

```

```

Alignment Scores:
Pred. No.:      61.4      Length:      804
Score:          40.00     Matches:      8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:      87.0% Indels:      0
DB:              10      Gaps:      0

```

```

US-10-774-176-11 (1-9) x BV564129 (1-804)

```

```

QY      2  LeuThrGluValProThrAspLeu 9
|||||
DB      666  CTACGGAGGTGCTACGGACCTC 689

```

```

RESULT 30
AF039719
LOCUS      35413 bp      DNA      linear      INV 21-SEP-2004
DEFINITION      Caenorhabditis elegans cosmid K04F10, complete sequence.
ACCESSION      AF039719
VERSION      AF039719.1 GI:2773236
KEYWORDS      HTG.
SOURCE      Caenorhabditis elegans
ORGANISM      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.

```

```

REFERENCE
AUTHORS      WormBase Consortium
CONSTRM      Genome sequence of the nematode C. elegans: a platform for
TITLE      investigating biology. The C. elegans Sequencing Consortium
JOURNAL      Science 282 (5396), 2012-2018 (1998)
PUBMED      9851916

```

```

REFERENCE
AUTHORS      Latreille,P. and Wamsley,P.
TITLE      The sequence of C. elegans cosmid K04F10
JOURNAL      Unpublished (2001)
AUTHORS      Waterston,R.
CONSTRM      3 (bases 1 to 35413)
TITLE      Direct Submission
JOURNAL      Submitted (06-JUL-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

```

```

REFERENCE
AUTHORS      Waterston,R.
TITLE      Direct Submission
JOURNAL      Submitted (01-AUG-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

```

```

REFERENCE
AUTHORS      Waterston,R.
TITLE      Direct Submission
JOURNAL      Submitted (23-MAY-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

```

```

REFERENCE
AUTHORS      Waterston,R.
TITLE      Direct Submission
JOURNAL      Submitted (19-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

```

```

REFERENCE
AUTHORS      Waterston,R.
TITLE      Direct Submission
JOURNAL      Submitted (26-DEC-2002) Department of Genetics, Washington

```

```

University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
8 (bases 1 to 35413)
Wilson,R.

```

```

Direct Submission
Submitted (15-JUN-2003) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
9 (bases 1 to 35413)

```

```

WormBase Consortium

```

```

Direct Submission

```

```

Submitted (21-SEP-2004) Department of Genetics, Washington

```

```

University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

```

```

Submitted by:

```

```

Genome Sequencing Center

```

```

Department of Genetics, Washington University

```

```

St. Louis , MO 63110, USA, and

```

```

Sanger Centre, Hinxton Hall

```

```

Cambridge CB10 1RQ, England

```

```

email: submissions@wustl.edu and jes@sanger.ac.uk

```

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

For a graphical representation of this clone sequence and its analysis see:

<http://www.wormbase.org/db/seq/sequence?name=K04F10;class=Sequence>

NEIGHBORING CLONE INFORMATION

The 5' clone is F26B1, 200 bp overlap; the 3' clone is ZC308, 600 bp overlap. Actual start of this clone is at base position 23932 of F26B1; actual end is at 4388 of ZC308.

NOTES:

Coding sequences below are the result of integration and manual review of the following data : computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFeome cloning project (<http://wormfb.dfci.harvard.edu/>), similarity to other proteins from BlastX analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES

source

1. 35413
/organism="Caenorhabditis elegans"

/mol_type="genomic DNA"

/strain="Bristol N2"

/db_xref="taxon:6239"

/chromosome="I"

/clone="K04F10"

10045..10046

/note="SL1 trans-splice acceptor; see yk846a02.5"

10252..24273

/gene="bli-4"

/locus_tag="K04F10.4"

misc_feature

gene

CDS

```
join(10252..10332,10377..10664,10772..10915,10976..11181,
11237..11554,11600..11680,11790..11966,12018..12245,
12292..12401,12454..12660,13241..13371,21787..21859,
22715..22893,22945..23053,23674..23901,23955..24108,
24159..24273)
/genes="bli-4"
/locus_tag="K04F10.4c"
/standard_name="K04F10.4c"
/notes="contains similarity to Pfam domains PF00082
(Subtilase family), PF01483 (Proprotein convertase
P-domain); coded for by the following C. elegans cDNAs:
CK580710, OSTF0988G7.1, OSTF090A5.1, OSTF090B11.1,
OSTF153A12.1, OSTF153B12.1, Yk19E5.3, Yk19E5.5, Yk25B5.5,
Yk36E7.5, Yk38G2.5, Yk129E7.5, Yk160A2.3, Yk160A2.5,
Yk208A3.3, Yk208A3.5, Yk218A9.3, Yk219C7.3, Yk219C7.5,
Yk238A1.3, Yk238A1.5, Yk263A3.5, Yk315C1.3, Yk315C1.5,
Yk315C7.5, Yk331G6.5, Yk350A10.3, Yk350A10.5, Yk406D2.5,
Yk408C4.5, Yk408D9.3, Yk533B2.3, Yk533B2.5, Yk602E5.3,
Yk602E5.5, Yk623A7.3, Yk623A7.5, Yk626C4.5, Yk780B02.5,
Yk791A05.5, Yk846A02.5, Yk1084F02.3, Yk1084F02.5,
Yk1113H11.3, Yk1113H11.5, Yk1126A07.3, Yk1126A07.5,
Yk1327G12.3, Yk1327G12.5, L29440, CEBSG74F, CEBSG74R,
Yk1425E03.3, Yk1425E03.5, Yk1433C06.5, Yk1459E12.5,
Yk1687B11.3, Yk1687B11.5"
/codon_start=1
/product="Blistered cuticle protein 4, isoform d"
/protein_id="AAB96754.1"
/db_xref="GI:2773243"
/db_xref="WormBase:K04F10.4d"
/translation="MRISIGRIANQILAVLIAVAFTHEDSICDESIGACGERIHTVI
RLAKRDELRAIAADHMDVKGDFLDTHFYLYSHSETTRRRHRAQVRELDSHPAV
EWEYEQPKRRKRDYLLDNDVHNSPNFRSVNLNRDTRRAQVQPSREIPLSPF
PDPLYKQDWYLGAGVGYDMNVKQAWLQVAGRNVSILDDGIQRDHPDLAANYDP
LASTDINDHDDPTPONNGDNKHGTRCAGEVAALAGNOCQGVAFKAKIGVGRMLDG
AVDSVEAASLSLNQDHIDIYSASWGPEDDKGTDPGCLAREAFYRGIKNGRGKGN
IFWASNGSRQSDCSADGYTTSVYTLSSATYDNRHPWYLEECSPSSATYSSAD
KPLNPNFGRNGVRGVRMSFGYGLDGGALVMNAKTWKTVEPHQICTEYRLANPN
PREIVGRQLNFTLDVNGCESGTPVLYLHVQVHATVYRLKRGDLKTLFSPSGTRSV
LLPFPQDQFNANGFHKWPFVLSVOQMGDPRTGTLWLMVESVTTPAACTFDHTWTLIY
GTADPAQGDVPVYATPATSQGVLSRVHQTLSOVESAPISPPDLTSAGNCHDRCNGG
CTESSATSCFACKHLTQTLRNKGGSGFKVCQKCDYDYLDBKCKMCSHCCTCTKA
EVCETPGSLLLLIDVNMHPYDHGKCVESCPGLVADYESNLVQAKIKWRKLDGGDY
YINAVGKDCIDCSCECTCTAPMGSCCKSGKSGISGICRCPCCPGSGTSQWQCDC
SKPDLTLIDSNKSSGFLMFVIVSLIAACGICACCKCASKSETKSSNVVAPLAQYNA
TNGAINLGHTDDDDDEVFVNPQIV"
join(10252..10332,10377..10664,10772..10915,10976..11181,
11237..11554,11600..11680,11790..11966,12018..12245,
12292..12401,12454..12660,13241..13371,21787..21859,
22715..22893,22945..23053,23096..23247)
/genes="bli-4"
/locus_tag="K04F10.4c"
/standard_name="K04F10.4c"
/notes="contains similarity to Pfam domains PF00082
(Subtilase family), PF01483 (Proprotein convertase
P-domain); coded for by the following C. elegans cDNAs:
OSTF0988G7.1, OSTF090A5.1, OSTF090B11.1, OSTF153A12.1,
OSTF153B12.1, OSTF090A5.1, Yk19E5.3, Yk25B5.5, Yk36E7.5,
Yk38G2.5, Yk129E7.5, Yk160A2.5, Yk208A3.5, Yk219C7.5,
Yk238A1.5, Yk263A3.5, Yk315C1.5, Yk315C7.5, Yk331G6.5,
Yk350A10.5, Yk406D2.5, Yk408C4.5, Yk533B2.5, Yk602E5.5,
Yk623A7.5, Yk626C4.5, Yk780B02.5, Yk791A05.5, Yk846A02.5,
Yk1084F02.5, Yk1113H11.5, Yk1126A07.5, Yk1327G12.5,
Yk1425E03.5, Yk1433C06.5, Yk1459E12.5, Yk1687B11.5"
/codon_start=1
/product="Blistered cuticle protein 4, isoform c"
/protein_id="AAB96757.1"
/db_xref="GI:2773246"
/db_xref="WormBase:K04F10.4c"
/translation="MRISIGRIANQILAVLIAVAFTHEDSICDESIGACGERIHTVI
RLAKRDELRAIAADHMDVKGDFLDTHFYLYSHSETTRRRHRAVRELDSHPAV
EWEYEQPKRRKRDYLLDNDVHNSPNFRSVNLNRDTRRAQVQPSREIPLSPF
PDPLYKQDWYLGAGVGYDMNVKQAWLQVAGRNVSILDDGIQRDHPDLAANYDP
```

CDS

```
LASTDINDHDDPTPONNGDNKHGTRCAGEVAALAGNOCQGVAFKAKIGVGRMLDG
AVDSVEAASLSLNQDHIDIYSASWGPEDDKGTDPGCLAREAFYRGIKNGRGKGN
IFWASNGSRQSDCSADGYTTSVYTLSSATYDNRHPWYLEECSPSSATYSSAD
KPLNPNFGRNGVRGVRMSFGYGLDGGALVMNAKTWKTVEPHQICTEYRLANPN
PREIVGRQLNFTLDVNGCESGTPVLYLHVQVHATVYRLKRGDLKTLFSPSGTRSV
LASTDINDHDDPTPONNGDNKHGTRCAGEVAALAGNOCQGVAFKAKIGVGRMLDG
AVDSVEAASLSLNQDHIDIYSASWGPEDDKGTDPGCLAREAFYRGIKNGRGKGN
IFWASNGSRQSDCSADGYTTSVYTLSSATYDNRHPWYLEECSPSSATYSSAD
KPLNPNFGRNGVRGVRMSFGYGLDGGALVMNAKTWKTVEPHQICTEYRLANPN
PREIVGRQLNFTLDVNGCESGTPVLYLHVQVHATVYRLKRGDLKTLFSPSGTRSV

Alignment Scores:
Pred. No.: 2.29e+03 Length: 35413
Score: 40.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 87.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-11 (1-9) x AF039719 (1-35413)
Oy 1 AsnLeuThrGluValProThrAspLeu 9
|||||
Db 3887 ANTTTGACACCTGTTCTCAGAGATTG 3913
|||||

AC151952 106590 bp DNA linear HTG 28-JUL-2005
DEFINITION Pan troglodytes clone rp43-149k8, WORKING DRAFT SEQUENCE, 15
unorderd pieces.
ACCESSION AC151952
VERSION AC151952.13 GI:71361787
KEYWORDS HTG: HTGS PHASE1; HTGS DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 106590)
AUTHORS Xu, W., Hua, A. and Roe, B.A.
JOURNAL Pan troglodytes BAC Clone rp43-149k8
TITLE Unpublished
REFERENCE 2 (bases 1 to 106590)
AUTHORS Xu, W., Hua, A. and Roe, B.A.
JOURNAL Direct Submission
TITLE Direct Submission
JOURNAL Submitted (14-OCT-2004) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Farrington Oval, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 106590)
AUTHORS Xu, W., Hua, A. and Roe, B.A.
JOURNAL Direct Submission
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2005) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2128: contig of 2128 bp in length
* * 2129 2228: gap of unknown length
* * 2229 4386: contig of 2158 bp in length
* * 4387 4486: gap of unknown length
* * 4487 6694: contig of 2208 bp in length
* * 6695 6794: gap of unknown length
* * 6795 8829: contig of 2035 bp in length
* * 8830 8929: gap of unknown length
* * 8930 11542: contig of 2613 bp in length
```


* 11543 11642: gap of unknown length
 * 11543 14864: contig of 3222 bp in length
 * 14865 14864: gap of unknown length
 * 14865 20301: contig of 5337 bp in length
 * 20302 20401: gap of unknown length
 * 20402 24081: contig of 3680 bp in length
 * 24082 24181: gap of unknown length
 * 24182 29193: contig of 5012 bp in length
 * 29194 29293: gap of unknown length
 * 29294 34714: contig of 5421 bp in length
 * 34715 34814: gap of unknown length
 * 34815 39112: contig of 4298 bp in length
 * 39113 39212: gap of unknown length
 * 39213 45068: contig of 5856 bp in length
 * 45069 45168: gap of unknown length
 * 45169 53899: contig of 8731 bp in length
 * 53900 53999: gap of unknown length
 * 54000 62745: contig of 8746 bp in length
 * 62746 106590: gap of unknown length
 * 62846 106590: contig of 43745 bp in length.

FEATURES

source

1. 106590
 Location/Qualifiers
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone_lib="rp43-149k8"
 /clone_lib="RP43-149k8"
 2129..2228

gap /estimated_length=unknown

gap 4387..4486

gap /estimated_length=unknown

gap 6695..6794

gap /estimated_length=unknown

gap 8330..8329

gap /estimated_length=unknown

gap 11543..11642

gap /estimated_length=unknown

gap 14865..14964

gap /estimated_length=unknown

gap 20302..20401

gap /estimated_length=unknown

gap 24082..24181

gap /estimated_length=unknown

gap 29194..29293

gap /estimated_length=unknown

gap 34715..34814

gap /estimated_length=unknown

gap 39113..39212

gap /estimated_length=unknown

gap 45069..45168

gap /estimated_length=unknown

gap 53900..53999

gap /estimated_length=unknown

gap 62746..62845

gap /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 6.57e+03 Length: 106590
 Score: 40.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 87.0% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-11 (1-9) x AC151952 (1-106590)

QY 2 LeuThrCluValProThrAspLeu 9

Db 15473 CTTACGGAGGTGCTACGGACCTC 15496

RESULT 32

AC166203

LOCUS AC166203 166980 bp DNA linear HTG 27-JUL-2005
 DEFINITION Oryctolagus cuniculus clone LBI-115F14, WORKING DRAFT SEQUENCE, 4
 unordered pieces.
 AC166203
 AC166203.1 GI:71274272
 VERSION HTG; HTGS PHASE1; HTGS DRAFT.
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
 SOURCE Oryctolagus cuniculus (rabbit)
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha;
 Leporidae; Oryctolagus.
 1 (bases 1 to 166980)
 Antonellis, A., Ayala, K., Bass, D., Benjamin, B., Bera, J., Chu, G.,
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G.,
 Coleman, H., Engle, J., Franks, S., Fukenko, T., Gestole, M.,
 Greene, A., Guan, X., Gupta, J., Gurson, N., Haghighi, P., Han, J.,
 Hansen, N., Ho, S.-L., Hu, P., Hunter, G., Hurle, B., Idol, J.R.,
 Kwong, P., Lalic, P., Larson, S., Lee-Lin, S.-Q., Legaspi, R.,
 Madden, M., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C.,
 Maskeri, B., McDowell, J., Mojidi, H.A., Mullikin, J.C., Park, M.,
 Portnoy, M.E., Prasad, A., Puri, O., Rantz, K., Reddix-Dugue, N.,
 Sante, A., Schandler, K., Schueler, M.G., Sison, C., Stantropop, S.,
 Tave, A., Thomas, J.W., Thomas, P.J., Tsipouri, V., Ung, L., Vogt, J.L.,
 Wetherby, K.D., Withers, T.R., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 166980)
 Green, E.D.
 Direct Submission
 Submitted (27-JUL-2005) NIH Intramural Sequencing Center, 5625
 Fishers Lane, Rockville, MD 20852, USA
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoengri.nih.gov
 ----- Project Information
 Center project name: llv
 Center clone name: 115F14
 ----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 16457 bases at least Q40
 Consensus quality: 165010 bases at least Q30
 Consensus quality: 165455 bases at least Q20
 Insert size: 168000; agarose-fp
 Insert size: 166880; sum-of-contigs
 Quality coverage: 9.04x in Q20 bases; agarose-fp
 Quality coverage: 9.11x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 11862: contig of 11862 bp in length
 * 11863 11962: gap of unknown length
 * 11963 23352: contig of 11390 bp in length
 * 23353 23452: gap of unknown length
 * 23453 87182: contig of 63730 bp in length
 * 87183 87282: gap of unknown length
 * 87283 166980: contig of 79698 bp in length.
 Location/Qualifiers
 1. 166980
 /organism="Oryctolagus cuniculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9986"
 /clone_lib="LBI-115F14"
 /clone_lib="LBI"

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1281: contig of 1281 bp in length
 * 1282 1381: gap of unknown length
 * 1382 7785: contig of 8404 bp in length
 * 7786 7885: gap of unknown length
 * 7886 25833: contig of 17948 bp in length
 * 25834 25933: gap of unknown length
 * 25934 53132: contig of 27199 bp in length
 * 53133 53232: gap of unknown length
 * 53233 102028: contig of 48796 bp in length
 * 102029 102128: gap of unknown length
 * 102129 181161: contig of 79033 bp in length.

FEATURES

source

1. 181161
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-285K4"
 /clone_lib="APCI human BAC library 11"
 1282..1381
 /estimated_length=unknown
 7786..7885
 /estimated_length=unknown
 25834..25933
 /estimated_length=unknown
 53133..53232
 /estimated_length=unknown
 102029..102128
 /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 1.09e+04 Length: 181161
 Score: 40.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 87.0% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-11 (1-9) x AC092347 (1-181161)

QY 2 LeuthrGluValProThrAspLeu 9

DB 8757 CTCACAGAGTCCCGGATCTC 8734

RESULT 35

AC166207/c

LOCUS AC166207 188254 bp DNA linear HTG 06-AUG-2005
 DEFINITION Oryctolagus cuniculus clone LB1-23912, WORKING DRAFT SEQUENCE, 5
 ordered pieces.

AC166207

AC166207.2 GI:71892504

HTG; HTGS_PHASE2; HTGS_DRAFT.

Oryctolagus cuniculus (rabbit)

SOURCE

Oryctolagus cuniculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha;

Leporidae; Oryctolagus.

1 (bases 1 to 188254)

Antoniellis,A., Ayele,K., Bass,D., Benjamin,B., Bera,J.,
 Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G.,
 Coleman,H., Engle,J., Franks,S., Fukukenko,T., Gestole,M.,
 Greene,A., Guan,X., Gupta,J., Gurson,N., Haghighi,P., Han,J.,
 Hansen,N., Ho,S.-L., Hunter,G., Hurler,B., Idol,J.R.,
 Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q., Legaspi,R.,
 Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
 Maskeri,B., McDowell,J., Mojidi,H.A., Mullikin,J.C., Park,M.,
 Portnoy,M.E., Prasad,A., Puri,O., Rantz,K., Reddix-Dugue,N.,
 Sante,A., Schandler,K., Schueler,M.G., Sison,C., Stantripop,S.,

REFERENCE

AUTHORS

FEATURES
 source

1. 188254
 /organism="Oryctolagus cuniculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9986"
 /clone="LB1-23912"
 /clone_lib="LB1"

Taye,A., Thomas,J.W., Thomas,P.J., Tsipouri,V., Ung,L., Vogt,J.L.,
 Wetherby,K.D., Withers,I.R., Young,A. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 188254)
 Green,E.D.
 Direct Submission
 Submitted (27-JUL-2005) NIH Intramural Sequencing Center, 5625
 Fishers Lane, Rockville, MD 20852, USA
 3 (bases 1 to 188254)
 Green,E.D.
 Direct Submission
 Submitted (06-AUG-2005) NIH Intramural Sequencing Center, 5625
 Fishers Lane, Rockville, MD 20852, USA
 On Aug 6, 2005 this sequence version replaced gi:71274277.

Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project Information

 Center project name: llw
 Center clone name: 239102

REFERENCE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is generally based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap, version 0.990319
 Consensus quality: 187477 bases at least Q40
 Consensus quality: 187640 bases at least Q30
 Consensus quality: 187774 bases at least Q20
 Insert size: 189000; agarose-fp
 Insert size: 187854; sum-of-contigs
 Quality coverage: 9.31x in Q20 bases; agarose-fp
 Quality coverage: 9.37x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 18794: contig of 18794 bp in length
 * 18795 18894: gap of unknown length
 * 18895 56668: contig of 37774 bp in length
 * 56669 56768: gap of unknown length
 * 56769 80395: contig of 23627 bp in length
 * 80396 80495: gap of unknown length
 * 80496 169236: contig of 8741 bp in length
 * 169237 169336: gap of unknown length
 * 169337 188254: contig of 18918 bp in length.

Location/Qualifiers

1. 188254
 /organism="Oryctolagus cuniculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9986"
 /clone="LB1-23912"
 /clone_lib="LB1"

/note="BAC resource: http://bacpac.chori.org/
breed: New Zealand White"

misc_feature

1..103190
/note="clone overlaps with GenBank Accession Number
AC166203 clone LB1-115F14 (center project name llv)"
1..18794
/note="assembly_fragment"

misc_feature

clone_end:SP6

gap

vector_side:left
18795..18894

misc_feature

/estimated_length=unknown
18895..56668

gap

56669..56768
/note="assembly_fragment"

misc_feature

/estimated_length=unknown
56769..80395

gap

80396..80495
/note="assembly_fragment"

misc_feature

80496..169236
/note="assembly_fragment"

gap

169237..169336
/estimated_length=unknown

misc_feature

169337..188254
/note="assembly_fragment"

misc_feature

clone_end:T7
vector_side:right

misc_feature

187910..188254
/note="clone overlaps with GenBank Accession Number
AC165396 clone LB1-64D22 (center project name llx)"

ORIGIN

Alignment Scores:
Pred. No.: 1.13e+04 Length: 188254
Score: 40.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 87.0% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-11 (1-9) x AC166207 (1-188254)

Qy 1 AsnLeuThrGluValProThrAspLeu 9

Db 6919 AATCTGGTTGAATATCCACAGACTA 6893

RESULT 36

AC146469/c AC146469 191684 bp DNA linear PRI 15-JUL-2004
LOCUS Pan troglodytes clone rp43-45g24, complete sequence.
DEFINITION AC146469

AC146469.23 GI:50300664

VERSION HTG.

KEYWORDS Pan troglodytes (chimpanzee)

SOURCE Pan troglodytes

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Pan.

REFERENCE 1 (bases 1 to 191684)

AUTHORS Xu, W. and Roe, B.A.

TITLE Pan troglodytes BAC Clone rp43-45g24

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 191684)

AUTHORS Xu, W. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (15-AUG-2003) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

REFERENCE 3 (bases 1 to 191684)

AUTHORS Xu, W. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (02-JUN-2004) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

4 (bases 1 to 191684)

AUTHORS Xu, W. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (15-JUL-2004) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

COMMENT On Jul 15, 2004 this sequence version replaced gi:47901752.

----- Genome Center

Center: Department Of Chemistry And Biochemistry

The University Of Oklahoma

Center code:UOKNOR

FEATURES

source

Location/Qualifiers

1..191684

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="rp43-45g24"

/clone_lib="RPCI - 43 Male Chimpanzee BAC Library"

ORIGIN

Alignment Scores:

Pred. No.: 1.15e+04 Length: 191684

Score: 40.00 Matches: 8

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 87.0% Indels: 0

DB: 8 Gaps: 0

US-10-774-176-11 (1-9) x AC146469 (1-191684)

Qy 2 LeuThrGluValProThrAspLeu 9

Db 47078 CTTACGGAGGTGCTACGACCTC 47055

RESULT 37

CR388209

LOCUS 231264 bp DNA linear HTG 11-AUG-2005

DEFINITION Danio rerio chromosome 16 clone DKEY-25C4, WORKING DRAFT SEQUENCE,

5 unordered pieces.

ACCESSION CR388209

VERSION CR388209.7 GI:72534292

KEYWORDS HTG; HTGS PHASE1; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 231264)

AUTHORS Phillimore, B.

TITLE Direct Submission

JOURNAL Submitted (10-AUG-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

zfish-help@sanger.ac.uk Clone requests:

http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeight

On Aug 12, 2005 this sequence version replaced gi:68161913.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: zfish-help@sanger.ac.uk

----- Project Information

Center project name: zK25C4

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 230238 bases at least Q40

Consensus quality: 230512 bases at least Q30

Consensus quality: 230578 bases at least Q20

Insert size: 230864; sum-of-contigs

Inert size: 223408; 2.8% error; agarose-fp

Quality coverage: 8.72x in Q20 bases; sum-of-contigs Quality

coverage: 9.70x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 130645: contig of 130645 bp in length
 130646 130745: gap of 100 bp
 130746 223377: contig of 92632 bp in length
 223378 223477: gap of 100 bp
 223478 226256: contig of 2779 bp in length
 226257 226356: gap of 100 bp
 226357 229095: contig of 2739 bp in length
 229096 229196: gap of 100 bp
 229196 231264: contig of 2069 bp in length.

FEATURES

source

1. .231264
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /chromosome="16"
 /clone="DKEY-25C4"
 /clone_lib="DanioKey"

misc_feature

1. .130645
 /note="assembly_fragment:01232
 fragment_chain:1"

misc_feature

130746. .223377
 /note="assembly_fragment:03846
 fragment_chain:1"

misc_feature

223478. .226256
 /note="assembly_fragment:00058"
 226357. .229095

misc_feature

/note="assembly_fragment:00100"
 229196. .231264
 /note="assembly_fragment:03982"

ORIGIN

Alignment Scores:
 Pred. No.: 1.38e+04 Length: 231264
 Score: 40.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 87.0% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-11 (1-9) x CR388209 (1-231264)

QY 2 LeuthrGluValProThrAspLeu 9

Db 107258 CTCACAGAAGTGCACACTGACCTA 107281

RESULT 38

BV090626

LOCUS 130746 394 bp DNA linear STS 15-OCT-2003
 DEFINITION BPMMSEQ0002525 Roche Palo Alto Mus musculus STS genomic, sequence tagged site.

ACCESSION

BV090626

VERSION

BV090626.1 GI:37668105

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Ueoka, J., Liao, G., Cheng, J., Nguyen, A., Bach, C., Puech, A.,

McPherson, J.D., Foernzler, D. and Peltz, G.

TITLE

Mus musculus SNPs

JOURNAL

Unpublished (2003)

COMMENT

Contact: Jonathan Usuka
 Roche Palo Alto Genetics and Genomics Department
 Roche Palo Alto
 3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
 Tel: 6508555807
 Email: Jonathan.Usuka@roche.com
 Primer A: No primer submitted
 Primer B: No primer submitted.

FEATURES

source

1. .394
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /map="11-18734-18346-AL672172.7.1.52478"
 /clone_lib="Roche Palo Alto"
 /note="SNPs developed from assay sequences derived from 15 different strains
 of mice (as of October 1, 2003). Those strains include
 A/J, A/HeJ,
 129/Sv, AKR/J, B10.D2-H2/cSnJ, BALB/cByJ, BALB/cJ,
 C3H/HeJ, C57BL/6J,
 CAST/Ei, DBA/2J, MRL/MpJ, NZB/BinJ, NZW/LaC, SPRET/Ei."
 <1. .394

STS

ORIGIN

Alignment Scores:
 Pred. No.: 51.5 Length: 394
 Score: 39.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 84.8% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-11 (1-9) x BV090626 (1-394)

QY 1 AsnLeuthrGluValProThrAspLeu 9

Db 151 ANCTTAGACAGATTCCTACAGATCTA 177

RESULT 39

AX525488/c

LOCUS

AX525488

DEFINITION

Sequence 10 from Patent WO02066682.

ACCESSION

AX525488

VERSION

AX525488.1 GI:25170366

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

Farris, G., Hicken, S.H. and Farr, S.B.

TITLE

Rat toxicologically relevant genes and uses thereof

JOURNAL

Patent: WO 02066682-A.10.29-AUG-2002;

Phase-1 Molecular Toxicology Inc. (US)

FEATURES

Location/Qualifiers

1. .691

/organism="Rattus norvegicus"

/mol_type="unassigned DNA"

/db_xref="taxon:10116"

ORIGIN

Alignment Scores:

Pred. No.: 88.1 Length: 691

Score: 39.00 Matches: 7

Percent Similarity: 88.9% Conservative: 1

Best Local Similarity: 77.8% Mismatches: 1

Query Match: 84.8% Indels: 0

DB: 6 Gaps: 0

```

US-10-774-176-11 (1-9) x AX525488 (1-691)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 430 AATCTTAGAGAGATTCTCAGATCTT 404

RESULT 40
AX438274
LOCUS AX438274 933 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 6689 from Patent WO0229113.
ACCESSION AX438274
VERSION AX438274.1 GI:21663082
KEYWORDS
SOURCE
ORGANISM Bacillus clausii
Bacillus clausii
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
AUTHORS Berka, R. and Clausen, I. G.
TITLE Methode for monitoring multiple gene expression
JOURNAL Patent: WO 0229113-A 6689 11-APR-2002;
Novozymes Biotech, Inc. (US) ; Novozymes A/S (DK)
FEATURES
source
location/Qualifiers
1..933
/organism="Bacillus clausii"
/mol_type="unassigned DNA"
/db_xref="taxon:79880"

ORIGIN
Alignment Scores:
Pred. No.: 117 Length: 933
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-11 (1-9) x AX438274 (1-933)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 669 AACTTGACTGAATGTCCTCACTGACATC 695

RESULT 41
AJ720394/C
LOCUS AJ720394 2949 bp mRNA linear VRT 12-JAN-2005
DEFINITION Gallus gallus mRNA for hypothetical protein, clone 16m20.
ACCESSION AJ720394
VERSION AJ720394.1 GI:53133447
KEYWORDS ORF1.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
AUTHORS Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezzubov, Y., Zaim, J.,
Fiedler, P., Kutter, S., Blagodatki, A., Kostovska, D., Koter, M.,
Plachy, J., Carninci, P., Hayashizaki, Y. and Buerstedde, J.M.
TITLE Full-length cDNAs from chicken bursal lymphocytes to facilitate
gene function analysis
JOURNAL Genome Biol. 6 (1), R6 (2005)
PUBMED 15642098
REFERENCE
AUTHORS Caldwell, R.B.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuherberg, GERMANY
FEATURES
source
location/Qualifiers
1..2949
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"

US-10-774-176-11 (1-9) x AX525488 (1-691)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 430 AATCTTAGAGAGATTCTCAGATCTT 404

RESULT 42
AX401746
LOCUS AX401746 2977 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 1422 from Patent WO0210453.
ACCESSION AX401746
VERSION AX401746.1 GI:21337926
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
Elashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 1422 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
source
location/Qualifiers
1..2977
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. M73714"

ORIGIN
Alignment Scores:
Pred. No.: 353 Length: 2949
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-11 (1-9) x AJ720394 (1-2949)
Qy 1 AsnLeuThrGluValProThrAspLeu 9
Db 1171 AACTTGACCCACGTGCCACACACATG 1145

RESULT 42
AX401746
LOCUS AX401746 2977 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 1422 from Patent WO0210453.
ACCESSION AX401746
VERSION AX401746.1 GI:21337926
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
Elashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 1422 07-FEB-2002;
Gene Logic, Inc. (US)
FEATURES
source
location/Qualifiers
1..2977
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. M73714"

ORIGIN
Alignment Scores:
Pred. No.: 356 Length: 2977
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1

```

```

Best Local Similarity: 77.8%  Mismatches: 1
Query Match: 84.8%  Indels: 0
DB: 6  Gaps: 0

US-10-774-176-11 (1-9) x AX401746 (1-2977)

QY 1 AsnLeuThrGluValProThrAspLeu 9
   |||||  |||||  |||||  |||||  |||||
Db 2373 ANTCITTAGAGAGATTCCTACAGATCTT 2399

RESULT 43
RATMAD
LOCUS
DEFINITION Rat microsomal aldehyde dehydrogenase mRNA, complete cds.
ACCESSION M73714
VERSION M73714.1 GI:205265
KEYWORDS aldehyde dehydrogenase, integral membrane protein.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2977)
AUTHORS Miyachi, K., Masaki, R., Taketani, S., Yamamoto, A., Akayama, M. and
          Tashiro, Y.
TITLE Molecular cloning, sequencing, and expression of cDNA for rat liver
          microsomal aldehyde dehydrogenase
JOURNAL J. Biol. Chem. 266 (29), 19536-19542 (1991)
PUBMED 1717467
COMMENT Original source text: Rattus norvegicus (strain Sprague-Dawley)
          (library: lambda-gt10; lambda-gt11) liver cDNA to mRNA.
          Location/Qualifiers
FEATURES
    source
        1..2977
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /tissue_type="liver"
            /tissue_lib="lambda-gt10; lambda-gt11"
    gene
        1..2977
            /gene="aldehyde dehydrogenase"
    CDS
        124..1578
            /gene="aldehyde dehydrogenase"
            /EC_number="1.2.1.3"
            /EC_number="1.2.1.5"
            /codon_start=1
            /product="aldehyde dehydrogenase"
            /protein_id="AAA41555.1"
            /db_xref="GI:205265"
            /translation="MERQVQLRQTFRSGSRPLRFLQQLALRRMVQREKDIILAA
            IADLSKSELNATSHVITILGIDFMGLNPLGASRPANCKNLLTMDEAYVQPEPL
            GVLLTGANNYPVLVLPVGAIAAGNAIVPSELSNTAKILAEPLQYLDQDLY
            MIVNGVETTRILRQRFHLYTGNTAVGKIIVMEAAKHLTPVTLGLGKSPCYIDR
            DCDDLVACRRITWGMNCGQTCIAPDYILCEASSODQIVOKIKDTVKDFYGNVKAS
            PDERLINLHFKRISLLEGGKIAQGGTDEATRIAPILTDVDPNSKVMQBEIFG
            PLIPVSVNVERAINFINDREKPLALYIFSHNNKLIKRVIDTSGGVGTGNDVIMHP
            TVNSLPFGVGAGSMGAYHGKYSFTFSHQRPCLLGLKGESVNNKLRYPNPSSEKSVSW
            SKFFLLKQFNKGRLQLLLVCLVAVAIVTKQL"

ORIGIN
Alignment Scores: 356 Length: 2977
Pred. No.: 39.00 Matches: 7
Score: 39.00
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-11 (1-9) x RATMAD (1-2977)

QY 1 AsnLeuThrGluValProThrAspLeu 9
   |||||  |||||  |||||  |||||  |||||
Db 2373 ANTCITTAGAGAGATTCCTACAGATCTT 2399

```

RESULT 44
BC003797
LOCUS

DEFINITION

Mus musculus aldehyde dehydrogenase family 3, subfamily A2, mRNA
(cDNA clone MGC:6055 IMAGE:3490519), complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 3007)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullihy, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smailus, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 3007)

Strausberg, R.

Direct Submission

Submitted (28-FEB-2001)

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 8 Row: m Column: 5

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 6680677.

Location/Qualifiers

1..3007

/organism="Mus musculus"

/mol_type="mRNA"

/strain="mix FVB/N_C57BL/6J"

/db_xref="taxon:10090"

/clone="MGC:6055 IMAGE:3490519"

/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months

old, gross tissue."

/clone_lib="NCI CGAP_Mam5"

/lab_host="DH10B"

FEATURES

source

```

gene
1. .3007
/name="Vector: pCMV-SPORT6"
/gene="Aldh3a2"
/db_xref="GeneID:11671"
/db_xref="MGI:1353452"
84. .1538
/gene="Aldh3a2"
/codon_start=1
/product="Aldh3a2 protein"
/protein_id="AAH03797.1"
/db_xref="GI:13277828"
/db_xref="GeneID:11671"
/db_xref="MGI:1353452"
/translation="MERQVLRQAQFSGRSRLRFLRQLEALRMVQREKEILAA
IAADLSKSELNAYSHEVITILGSDFMGLNPGLASPAKKNLLTMMDSAYVQPEPL
GVLLIIGAWNPVFLTMQPLVGAIAAGNAAIKVPSELSSENTAKIELLFQYLDQDLY
AIVNGIGIPETELLKORFHLITGTAVGKI VMEAAAKHLTPVLELGGKSPCYIDR
DCDLVACRIAMGKYNCGTCTIADPYIICBASLQNIYVKIKETVKDFVGENIKAS
PDYERLINLHFPRLQSLKAGKIAEGGEMDEATRYLAPTILIDVDNSKVMQBEIFG
PILPVSVKQVDEAIPINDREKPLALYVSRNNKLIRVIDETSSGGVTGNDVIMHF
TVNSLPFGVGASGMGAYHGKYSFDTFSHQRPCLLGLKRGESVNNKLRYPNPSSEKVSWM
AKFFLLKQFNKGRGLGMLLFVCLVAVAIVKQQL"
84. .1340
/misc_feature
/gene="Aldh3a2"
/note="Aldh3a2; Region: Aldehyde dehydrogenase family. This
family of dehydrogenases act on aldehyde substrates.
Members use NADP as a cofactor. The family includes the
following members: The prototypical members are the
aldehyde dehydrogenases EC:1.2.1.3. Succinate-semialdehyde
dehydrogenase EC:1.2.1.16. Lactaldehyde dehydrogenase
EC:1.2.1.22. Benzaldehyde dehydrogenase EC:1.2.1.28.
Methylmalonate-semialdehyde dehydrogenase EC:1.2.1.27.
Glyoxaldehyde-3-phosphate dehydrogenase EC:1.2.1.9.
Delta-1-pyrroline-5-carboxylate dehydrogenase EC:
1.5.1.12. Acetaldehyde dehydrogenase EC:1.2.1.10.
Glutamate-5-semialdehyde dehydrogenase EC:1.2.1.41. This
family also includes omega crystallin, an eye lens protein
from squid and octopus that has little aldehyde
dehydrogenase activity"
/db_xref="CDD:pfam00171"

ORIGIN
Alignment Scores:
Pred. No.: 360 Length: 3007
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-11 (1-9) x BC003797 (1-3007)

QY 1 AsnLeuThrGluValProThrAspLeu 9
||||| ||||| ||||| ||||| |||||
2377 AATCTTAGAGAGATTCCTACAGATCTA 2403

RESULT 45
AP004979 30916 bp DNA linear PLN 22-JUL-2003
LOCUS
DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 4,
clone:LjT44L05, TW0162b, complete sequence.
ACCESSION AP004979.1 GI:21907997
VERSION
KEYWORDS HTG.
SOURCE
ORGANISM Lotus corniculatus var. japonicus (Lotus japonicus)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
REFERENCE
1
Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S.
Structural Analysis of a Lotus japonicus Genome. I. Sequence

Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
Regions of the Genome
Unpublished
2 (bases 1 to 30916)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL
Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:yn@kazusa.or.jp,
URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,
Fax:81-438-52-3934)
FEATURES
Location/Qualifiers
source
1..30916
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="4"
/clone="LjT44L05"
/clone_lib="LjT library"
/note="TW0162b, a part of TAC clone:TW0162.
synonym: Lotus japonicus"

ORIGIN
Alignment Scores:
Pred. No.: 3.34e+03 Length: 30916
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 15 Gaps: 0

US-10-774-176-11 (1-9) x AP004979 (1-30916)

QY 1 AsnLeuThrGluValProThrAspLeu 9
||||| ||||| ||||| ||||| |||||
Db 12691 AATCTTAGAGATTCCTACAGATCTC 12665

RESULT 46
AL672172/c
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-133K20 on chromosome 11 Contains
the Aldh3a2 gene for aldehyde dehydrogenase family 3 subfamily A2
and a CpG island, complete sequence.
ACCESSION AL672172
VERSION AL672172.7 GI:21425250
KEYWORDS HTG; aldehyde dehydrogenase; Aldh3a2; CpG island.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 52478)
Tromans,A.
Direct Submission
Submitted (04-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jun 13, 2002 this sequence version replaced gi:20803600.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
-----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

```


30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

RP23-133K20 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6.

FEATURES

source

Location/Qualifiers

1..52478
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-133K20"
 /clone_lib="RPCI-23"
 complement(join(17971..19439,21980..22215,24061..24160,26742..26908,29857..29998,31859..31976,35337..35545,37364..37449,38182..38413,39972..40303))
 /gene="Aldh3a2"

gene

/locus_tag="RP23-5512.7-002"
 complement(join(17971..19439,21980..22215,24061..24160,26742..26908,29857..29998,31859..31976,35337..35545,37364..37449,38182..38413,39972..40303))
 /gene="Aldh3a2"

mRNA

/locus_tag="RP23-5512.7-002"
 complement(join(17971..19439,21980..22215,24061..24160,26742..26908,29857..29998,31859..31976,35337..35545,37364..37449,38182..38413,39972..40303))
 /gene="Aldh3a2"
 /product="aldehyde dehydrogenase family 3, subfamily A2"
 /notes="match: ESTs: AI255724.1 AJ785140.1 BB848321.1 BF180302.1 BF784089.1 BI851844.1 BQ179769.1 BQ712918.1 BQ330804.1 CB955482.1
 match: cDNAs: AK079571.1 AK079639.1 AK086262.1 BC003797.1 M73714.1"
 complement(17971)
 /gene="Aldh3a2"

polya_site

/locus_tag="RP23-5512.7-002"
 complement(17984..17989)
 /gene="Aldh3a2"

polya_signal

/locus_tag="RP23-5512.7-002"
 complement(join(19428..19439,21980..22215,24061..24160,26742..26908,29857..29998,31859..31976,35337..35545,37364..37449,38182..38413,39972..40124))
 /gene="Aldh3a2"

CDS

/locus_tag="RP23-5512.7-002"
 /standard_name="OTTMUSP00000006018"
 /notes="match: proteins: P47740 Q99L64 Q99PH3"
 /codon_start=1
 /product="aldehyde dehydrogenase family 3, subfamily A2"
 /protein_id="CAI24063.1"
 /db_xref="GI:56206600"
 /db_xref="InterPro:IPR002086"
 /translation="MERQVLRQAFRSGRSRLRFLRQLQLEALRRMVQREKEILAA
 IAADLSKSELNAYSHVEITILGEIDFNLGNLPELASARPAKKNLLTWMDEAYVQPEPL
 GVLLIGAWNYPFVLTWQPLVGAAGNAIVKPSSELSSENTAKILAEILLQYLDQDLY
 AIVNGGIPETTELKQRFHLYTGTAVGKI VMEAAKHLTPVTLELGGSPCYIDR
 DCDDLVACRIANGKYNCCQTCAIPDYILCEASLQNIQVKIKETVKDFYGENIKAS
 PDYERITNLRHFKRLQSLKQKIAFGEMDEATRYLAPTILTDDVPSNKVMQBEIFG
 PILPIVSKNVDAINFINDREKPLALYVFSRNKKIKRVIDETSSGGVTGNDVIMHF
 TVNSLPGGVGASGCMGAYHGKYSPTFSHORPCLLKGLKGSVNKLRYPPNSSEKYSW
 AKFPLLKQFNKRGMLLFFVCLVAVAIVYKFLERA"
 complement(17984..17989)
 /gene="Aldh3a2"

gene

/locus_tag="RP23-5512.7-003"
 complement(join(20645..21600,21980..22215,24061..24160,26742..26908,29857..29998,31859..31976,35337..35545,37364..37449,38182..38413,39972..40303))
 /gene="Aldh3a2"

mRNA

/locus_tag="RP23-5512.7-003"
 complement(join(20645..21600,21980..22215,24061..24160,26742..26908,29857..29998,31859..31976,35337..35545,37364..37449,38182..38413,39972..40303))
 /gene="Aldh3a2"

CDS

/product="aldehyde dehydrogenase family 3, subfamily A2"
 /notes="match: ESTs: BB848321.1 BU963092.1
 match: cDNAs: AK051219.1"
 complement(join(21580..21600,21980..22215,24061..24160,26742..26908,29857..29998,31859..31976,35337..35545,37364..37449,38182..38413,39972..40124))
 /gene="Aldh3a2"
 /locus_tag="RP23-5512.7-003"
 /standard_name="OTTMUSP00000006019"
 /notes="match: proteins: Q8BQ93 Q8C3C9"
 /codon_start=1
 /product="aldehyde dehydrogenase family 3, subfamily A2"
 /protein_id="CAI24064.1"
 /db_xref="GI:56206601"
 /translation="MERQVLRQAFRSGRSRLRFLRQLQLEALRRMVQREKEILAA
 IAADLSKSELNAYSHVEITILGEIDFNLGNLPELASARPAKKNLLTWMDEAYVQPEPL
 GVLLIGAWNYPFVLTWQPLVGAAGNAIVKPSSELSSENTAKILAEILLQYLDQDLY
 AIVNGGIPETTELKQRFHLYTGTAVGKI VMEAAKHLTPVTLELGGSPCYIDR
 DCDDLVACRIANGKYNCCQTCAIPDYILCEASLQNIQVKIKETVKDFYGENIKAS
 PDYERITNLRHFKRLQSLKQKIAFGEMDEATRYLAPTILTDDVPSNKVMQBEIFG
 PILPIVSKNVDAINFINDREKPLALYVFSRNKKIKRVIDETSSGGVTGNDVIMHF
 TVNSLPGGVGASGCMGAYHGKYSPTFSHORPCLLKGLKGSVNKLRYPPNSSEKYSW
 AKFPLLKQFNKRGMLLFFVCLVAVAIVYKFLERA"
 join(complement(39972..40303), complement(38182..38413), complement(37364..37449), complement(35337..35545), complement(31859..31976), complement(29857..29998), complement(26742..26908), complement(24061..24160), complement(21980..22215), complement(1646093..9:235935..237204))
 /gene="Aldh3a2"
 /locus_tag="RP23-5512.7-001"
 join(complement(39972..40303), complement(38182..38413), complement(37364..37449), complement(35337..35545), complement(31859..31976), complement(29857..29998), complement(26742..26908), complement(24061..24160), complement(21980..22215), complement(1646093..9:235935..237204))
 /gene="Aldh3a2"

mRNA

/locus_tag="RP23-5512.7-001"
 /product="aldehyde dehydrogenase family 3, subfamily A2"
 /notes="match: ESTs: AV343011.1 BB848321.1 BF1551573.1 BG086218.1 BI144185.1 BI150323.1 BQ928400.1 BY745351.1
 match: cDNAs: AK078243.1"
 join(complement(39972..40124), complement(38182..38413), complement(37364..37449), complement(35337..35545), complement(31859..31976), complement(29857..29998), complement(26742..26908), complement(24061..24160), complement(21980..22215), complement(1646093..9:237049..237204))
 /gene="Aldh3a2"

CDS

/locus_tag="RP23-5512.7-001"
 /standard_name="OTTMUSP00000006017"
 /codon_start=1
 /product="aldehyde dehydrogenase family 3, subfamily A2"
 /protein_id="CAI24065.1"
 /db_xref="GI:56206602"
 /translation="MERQVLRQAFRSGRSRLRFLRQLQLEALRRMVQREKEILAA
 IAADLSKSELNAYSHVEITILGEIDFNLGNLPELASARPAKKNLLTWMDEAYVQPEPL
 GVLLIGAWNYPFVLTWQPLVGAAGNAIVKPSSELSSENTAKILAEILLQYLDQDLY
 AIVNGGIPETTELKQRFHLYTGTAVGKI VMEAAKHLTPVTLELGGSPCYIDR
 DCDDLVACRIANGKYNCCQTCAIPDYILCEASLQNIQVKIKETVKDFYGENIKAS
 PDYERITNLRHFKRLQSLKQKIAFGEMDEATRYLAPTILTDDVPSNKVMQBEIFG
 PILPIVSKNVDAINFINDREKPLALYVFSRNKKIKRVIDETSSGGVTGNDVIMHF
 TVNSLPGGVGASGCMGAYHGKYSPTFSHORPCLLKGLKGSVNKLRYPPNSSEKYSW
 AKFPLLKQFNKRGMLLFFVCLVAVAIVYKFLERA"
 complement(22190..22215), complement(24061..24160), complement(21980..22215), complement(1646093..9:237049..237204))
 /gene="Aldh3a2"

gene

/locus_tag="RP23-5512.7-001"
 /standard_name="OTTMUSP00000006017"
 /codon_start=1
 /product="aldehyde dehydrogenase family 3, subfamily A2"
 /protein_id="CAI24065.1"
 /db_xref="GI:56206602"
 /translation="MERQVLRQAFRSGRSRLRFLRQLQLEALRRMVQREKEILAA
 IAADLSKSELNAYSHVEITILGEIDFNLGNLPELASARPAKKNLLTWMDEAYVQPEPL
 GVLLIGAWNYPFVLTWQPLVGAAGNAIVKPSSELSSENTAKILAEILLQYLDQDLY
 AIVNGGIPETTELKQRFHLYTGTAVGKI VMEAAKHLTPVTLELGGSPCYIDR
 DCDDLVACRIANGKYNCCQTCAIPDYILCEASLQNIQVKIKETVKDFYGENIKAS
 PDYERITNLRHFKRLQSLKQKIAFGEMDEATRYLAPTILTDDVPSNKVMQBEIFG
 PILPIVSKNVDAINFINDREKPLALYVFSRNKKIKRVIDETSSGGVTGNDVIMHF
 TVNSLPGGVGASGCMGAYHGKYSPTFSHORPCLLKGLKGSVNKLRYPPNSSEKYSW
 AKFPLLKQFNKRGMLLFFVCLVAVAIVYKFLERA"
 complement(22190..22215,24061..24160,26742..26870,29857..29998,31859..31976,35337..35453))
 /gene="Aldh3a2"
 /locus_tag="RP23-5512.7-007"
 complement(join(22190..22215,24061..24160,26742..26870,29857..29998,31859..31976,35337..35453))
 /gene="Aldh3a2"

mRNA

Db 979 TTACAGARATTCCTACTGACCTG 1002

RESULT 49

AP006627_40

WPCOMMENT

Sequence split into 43 fragments LOCUS AP006627 Accession AP006627

Fragment Name Begin End
AP006627_00 1 110000
AP006627_01 100001 210000
AP006627_02 200001 310000
AP006627_03 300001 410000
AP006627_04 400001 510000
AP006627_05 500001 610000
AP006627_06 600001 710000
AP006627_07 700001 810000
AP006627_08 800001 910000
AP006627_09 900001 1010000
AP006627_10 1000001 1110000
AP006627_11 1100001 1210000
AP006627_12 1200001 1310000
AP006627_13 1300001 1410000
AP006627_14 1400001 1510000
AP006627_15 1500001 1610000
AP006627_16 1600001 1710000
AP006627_17 1700001 1810000
AP006627_18 1800001 1910000
AP006627_19 1900001 2010000
AP006627_20 2000001 2110000
AP006627_21 2100001 2210000
AP006627_22 2200001 2310000
AP006627_23 2300001 2410000
AP006627_24 2400001 2510000
AP006627_25 2500001 2610000
AP006627_26 2600001 2710000
AP006627_27 2700001 2810000
AP006627_28 2800001 2910000
AP006627_29 2900001 3010000
AP006627_30 3000001 3110000
AP006627_31 3100001 3210000
AP006627_32 3200001 3310000
AP006627_33 3300001 3410000
AP006627_34 3400001 3510000
AP006627_35 3500001 3610000
AP006627_36 3600001 3710000
AP006627_37 3700001 3810000
AP006627_38 3800001 3910000
AP006627_39 3900001 4010000
AP006627_40 4000001 4110000
AP006627_41 4100001 4210000
AP006627_42 4200001 4303871

Continuation (41 of 43) of AP006627 from base 4000001 (AP006627 Bacillus clausii KSM-K16

Alignment Scores:
Pred. No.: 1.12e+04 Length: 110000
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 1 Gaps: 0

US-10-774-176-11 (1-9) x AP006627_40 (1-110000)

Qy 1 AsnLeuThrGluValProThrAspLeu 9

Db 32372 AACTTGACTGAATGTCCCGACATC 32398
|||||

RESULT 50

AE017353_2/c

WPCOMMENT

Sequence split into 8 fragments LOCUS AE017353 Accession AE017353

Fragment Name Begin End

AE017353_0 1 110000
AE017353_1 100001 210000

AE017353_2 200001 310000
AE017353_3 300001 410000
AE017353_4 400001 510000
AE017353_5 500001 610000
AE017353_6 600001 710000
AE017353_7 700001 787999
Continuation (3 of 8) of AE017353 from base 200001 (AE017353 Cryptococcus neoformans var

Alignment Scores:
Pred. No.: 1.12e+04 Length: 110000
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 84.8% Indels: 0
DB: 15 Gaps: 0

US-10-774-176-11 (1-9) x AE017353_2 (1-110000)

Qy 1 AsnLeuThrGluValProThrAspLeu 9

Db 68561 AACATCGCAGGTACCGACGACCTC 68535
|||||

Search completed: April 25, 2006, 20:31:24
Job time : 3106.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds
(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-10

Perfect score: 40

Sequence: 1 FLGIVLALI 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-O=/abes/ABSSWEB spool/US10774176/runat_24042006_165112_19185/app_query.fasta.1
-DB=N Geneseg -OPWT=fastap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05p
-USER=US10774176 @CGN 1.1 3463 @runat_24042006_165112_19185 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 21.*

1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	40	100.0	108 10	ACD97670 Human col
2	40	100.0	475 13	ADU11677 Solid tum
3	40	100.0	927 6	ABT07721 Breast ca
4	40	100.0	927 8	ABX76333 Lung canc

5	40	100.0	927	10	ADB80503
6	40	100.0	927	11	ADN38723
7	40	100.0	973	8	AAV56198
8	40	100.0	1156	6	ABV99349
9	40	100.0	1260	6	ABK87175
10	40	100.0	1260	10	ADB97513
11	40	100.0	1260	10	ADB97452
12	40	100.0	1263	3	AAA27058
13	40	100.0	1263	4	AAF89736
14	40	100.0	1263	6	ABK87174
15	40	100.0	1281	3	AAA27059
16	40	100.0	1331	8	AAV56199
17	40	100.0	2020	10	ADJ56299
18	40	100.0	2053	8	ACC51052
19	40	100.0	2053	8	ABX76332
20	40	100.0	2053	8	AAV56197
21	40	100.0	2053	8	AAV56200
22	40	100.0	2053	11	ADN38721
23	40	100.0	2053	12	ADL06473
24	40	100.0	2053	12	ADN03961
25	40	100.0	2053	13	ADR25444
26	40	100.0	2053	13	ACN38510
27	40	100.0	2053	13	ADV35098
28	40	100.0	2338	5	AAV87175
29	40	100.0	2359	4	AAK94253
30	40	100.0	2359	12	ADL30831
31	40	100.0	2361	4	AAK94254
32	40	100.0	2361	12	ADL26162
33	40	100.0	2361	12	ADL30833
34	40	100.0	2557	12	ADL26160
35	40	100.0	2557	12	ADL26158
c 36	36	90.0	207	12	ADL03007
c 37	36	90.0	312	6	ABK79121
c 38	36	90.0	402	4	AAV37087
c 39	36	90.0	575	10	ADK53767
c 40	36	90.0	689	4	AAV19859
c 41	36	90.0	829	6	ABK79169
c 42	36	90.0	1833	8	ACA39024
c 43	36	90.0	2271	8	ACA21440
c 44	36	90.0	96109	4	AAF28548
c 45	36	90.0	168407	13	ABD33266
c 46	36	90.0	349980	6	ABQ81842
c 47	35	87.5	666	8	ACF73669
c 48	35	87.5	853	6	ABK63032
c 49	35	87.5	853	6	ABT11637
c 50	35	87.5	1326	3	ACV37238
c 51	35	87.5	1390	13	ADT15939
c 52	35	87.5	2485	4	AAK81357
c 53	35	87.5	3584	11	ADM02060
c 54	35	87.5	5738	2	AAV74395
c 55	35	87.5	11334	6	ABL33233
c 56	35	87.5	12301	13	ADR68371
c 57	35	87.5	16569	14	ADW94172
c 58	35	87.5	20467	4	AAV36361
c 59	35	87.5	20467	4	AAV36360
c 60	35	87.5	20467	8	ABX59348
c 61	35	87.5	20467	8	ABX59349
c 62	35	87.5	20467	12	ADJ30098
c 63	35	87.5	20467	12	ADJ30099
c 64	35	87.5	36221	10	ADH63075
c 65	35	87.5	110000	14	ABE42401
c 66	35	87.5	140167	6	ABT10146
c 67	34	85.0	418	9	ACH19908
c 68	34	85.0	422	10	ADB49910
c 69	34	85.0	422	13	ADV39405
c 70	34	85.0	426	8	ABZ55964
c 71	34	85.0	573	11	ACL34364
c 72	34	85.0	573	12	ADJ44733
c 73	34	85.0	588	13	ACN53260
c 74	34	85.0	630	6	ABK63060
c 75	34	85.0	630	10	ADB51738
c 76	34	85.0	630	10	ABT41392
c 77	34	85.0	630	12	ADP72253

ADB80503	Ovarian c
Adn38723	Cancer/an
Aad56198	Human LRR
Abv99349	Human NOV
Abk87175	cdNA enco
Adb97513	Feline 5T
Adb97452	DNA enco
Aaa27058	Human 5T4
Aaf89736	Nucleotid
Abk87174	cdNA enco
Aaa27059	Mouse 5T4
Aad56199	Human LRR
Adj56299	Human cDN
Acc51052	Human bla
Abx76332	Lung canc
Aad56197	Human LRR
Aad56200	Human LRR
Adn38721	Cancer/an
Adl06473	Human tum
Adn03961	Antipeori
Adr25444	Breast ca
Adr38510	Tumour-as
Adv35098	Human cDN
Aas87175	DNA enco
Aak94253	Human ful
Adl30831	Full leng
Aak94254	Human ful
Adl26162	Human cDN
Adl30833	Full leng
Adl26160	Human cDN
Adl26158	Human cDN
Adl03007	DNA enco
Abk79121	Bacillus
Aas37087	Novel hum
Adk53767	Plant DNA
Aal19859	Human bre
Abk79169	Bacillus
Acas9024	Prokaryot
Acas21440	Prokaryot
Aaf28548	Genomic f
Abd33266	Murine ca
Abq81842	Bifidobac
Acf73669	Staphyloc
Abk63032	Selected
Abt11637	Yeast sel
Aac37238	Arabidops
Adt15939	Plant cDN
Adm02060	Human cDN
Aav74395	Staphyloc
Abi33233	Human imm
Adt68371	Swine fev
Adw94172	Staphyloc
Aal36361	Human mus
Aal36360	Human mus
Abx59348	cdNA enco
Abx59349	cdNA enco
Adj30098	Human mus
Adj30099	Human mus
Adh63075	Human fib
Abt10146	Human bre
Abt10146	Human bre
Adb49910	Primary r
Adv39405	Rat cardi
Abz55964	Aspergill
ACL34364	Rice abio
Adj44733	Plant cDN
Acn53260	Cotton an
Abk63060	Rat seque
Adb51738	Primary r
Abt41392	Toxicity
Adp72253	Renal tox

78	34	85.0	727	13	ADK29620	Adx29620 Plant ful	151	33	82.5	1213	6	ABZ15649	Abz15649 Arabidops
79	34	85.0	752	13	ADK1268	Adk1268 Cotton cd	152	33	82.5	1251	6	ABZ67314	Abz67314 Breast sp
80	34	85.0	841	8	ABX12768	Abx12768 DNA encod	c 153	33	82.5	1302	6	AAD29698	Aad29698 Bacteriop
81	34	85.0	1149	6	ABA90338	Abag90338 Human pol	c 154	33	82.5	1305	10	ADC90794	Adc90794 E. faeciu
82	34	85.0	1284	4	ABL06047	Ablo6047 Drosophil	c 155	33	82.5	1320	6	AAD29697	Aad29697 Bacteriop
83	34	85.0	1413	11	ABD00115	Abd00115 Klebsiell	156	33	82.5	1320	10	ADC90887	Adc90887 E. faeciu
84	34	85.0	1500	11	ACH99827	Ach99827 Klebsiell	157	33	82.5	1350	8	ACA43823	Aca43823 Prokaryot
85	34	85.0	1772	5	AAS86183	Aas86183 DNA encod	158	33	82.5	1359	6	ABN68580	Abn68580 Streptoco
86	34	85.0	1810	2	AQ14978	Aq14978 Acya gene	159	33	82.5	1365	4	AAS52997	Aas52997 Enterococ
87	34	85.0	1857	12	ADK16386	Adk16386 Nanoarcha	160	33	82.5	1368	6	ABN68579	Abn68579 Streptoco
88	34	85.0	2212	10	ADA52877	Ada52877 Human cod	161	33	82.5	1371	6	ABN70478	Abn70478 Streptoco
89	34	85.0	2446	8	ADA63603	Ada63603 Maize sta	162	33	82.5	1371	13	ADV84221	Adv84221 Streptoco
90	34	85.0	2526	10	ADA69656	Ada69656 Rice gene	c 163	33	82.5	1374	13	ADK60350	Adk60350 Bacterial
91	34	85.0	2848	4	ABL17915	Abli17915 Drosophil	164	33	82.5	1400	6	ABQ68891	Abq68891 Listeria
92	34	85.0	3530	13	ADR08289	Adr08289 Full leng	c 165	33	82.5	1548	8	ACA48039	Aca48039 Prokaryot
93	34	85.0	3993	13	ADU01816	Adu01816 Novel hum	c 166	33	82.5	1589	12	ADP89729	Adp89729 Human Clo
94	34	85.0	4315	4	ABL06046	Abli06046 Drosophil	c 167	33	82.5	1593	13	AAA54296	Aaa54296 ORF of co
95	34	85.0	6131	4	ABL17914	Abli17914 Drosophil	c 168	33	82.5	1595	12	ADP89735	Adp89735 Human Clo
96	34	85.0	6247	6	ABL32275	Abli32275 Human imm	c 169	33	82.5	1595	12	ADP89730	Adp89730 Human Clo
97	34	85.0	10996	4	AAS46806	Aas46806 Tumour su	c 170	33	82.5	1595	12	ADP89733	Adp89733 Human Clo
98	34	85.0	10996	6	ABK28466	Abk28466 DNA trans	c 171	33	82.5	1595	12	ADP89737	Adp89737 Human Clo
99	34	85.0	28432	4	ABL05010	Abli05010 Drosophil	c 172	33	82.5	1596	12	ADP89734	Adp89734 Human Clo
100	34	85.0	36135	4	AAK84218	Aak84218 Human imm	c 173	33	82.5	1596	12	ADP89731	Adp89731 Human Clo
101	34	85.0	110000	2	AAK91990	Continuation (12 o	c 174	33	82.5	1596	12	ADP89727	Adp89727 Human Clo
102	34	85.0	110000	12	ADK16049	Continuation (2 of	c 175	33	82.5	1596	12	ADP89726	Adp89726 Human Clo
103	34	85.0	115284	11	ACN44296	Acn44296 Mouse gen	c 176	33	82.5	1596	12	ADP89736	Adp89736 Human Clo
104	34	85.0	128034	10	ADR43582	Adr43582 Polymorph	c 177	33	82.5	1630	12	ADG32103	Adg32103 DNA encod
105	34	85.0	128034	10	ADR43581	Adr43581 Human IDE	c 178	33	82.5	1637	12	ADP89728	Adp89728 Human Clo
106	34	85.0	128034	12	ADH54059	Adh54059 Human IDE	c 179	33	82.5	1723	13	ADK15393	Adk15393 Plant ful
107	34	85.0	128034	12	ADH54060	Adh54060 Human IDE	c 180	33	82.5	1732	4	ABL15627	Abli15627 Drosophil
108	34	85.0	174600	12	ADQ97520	Adq97520 Mouse can	c 181	33	82.5	1767	13	ADK14869	Adk14869 Plant ful
109	34	85.0	202100	10	ADR43315	Adr43315 Human IDE	c 182	33	82.5	1775	12	ADP89732	Adp89732 Human Clo
110	34	85.0	202100	12	ADH54357	Adh54357 Human IDE	c 183	33	82.5	1815	13	ADT44182	Adt44182 Bacterial
111	34	85.0	269223	4	AAK78554	Aak78554 Genomic f	c 184	33	82.5	1872	6	ABZ32532	Abz32532 Candida a
112	34	85.0	349960	6	ABQ81847	Abq81847 Bifidobac	c 185	33	82.5	1878	1	AAK90416	Aak90416 DNA seque
113	34	85.0	349960	6	ABQ81848	Abq81848 Bifidobac	c 186	33	82.5	2000	8	ADA71798	Ada71798 Rice gene
114	33	82.5	89	2	AAV76174	Aav76174 Staphyloc	c 187	33	82.5	2005	2	AAK13406	Aak13406 Enterococ
115	33	82.5	105	8	ACP74724	AcP74724 Staphyloc	c 188	33	82.5	2005	6	ABG99201	Abg99201 Enterococ
116	33	82.5	198	6	ABN92026	Abn92026 Staphyloc	c 189	33	82.5	2143	3	AAF16116	Aaf16116 Human pro
117	33	82.5	267	6	AAO11198	Aao11198 Human rep	c 190	33	82.5	2143	6	ABL89884	Abli89884 Human pol
118	33	82.5	267	4	ABL96657	Abli96657 Human tes	c 191	33	82.5	2225	6	ABQ70731	Abq70731 Listeria
119	33	82.5	303	3	AAK30790	Aak30790 Human sec	c 192	33	82.5	2245	14	ADY68649	Ady68649 Human v-c
120	33	82.5	351	3	AAA56854	Aaa56854 Mycoplasma	c 193	33	82.5	2250	11	ADM03387	Adm03387 Human cdn
121	33	82.5	357	6	ABL66969	Abli66969 Thyroid c	c 194	33	82.5	2318	12	ADJ75212	Adj75212 Marker ge
122	33	82.5	406	4	AAK12040	Aak12040 Human bre	c 195	33	82.5	2318	12	ADL63180	Adl63180 Human PRO
123	33	82.5	453	4	AAH53465	Aah53465 S. epider	c 196	33	82.5	2358	4	AAF26081	Aaf26081 FIV gp140
124	33	82.5	459	6	ABN92203	Abn92203 Staphyloc	c 197	33	82.5	2453	3	AAA54295	Aaa54295 Consensus
125	33	82.5	459	13	ADG01974	Adg01974 Staphyloc	c 198	33	82.5	2503	4	AAF59602	Aaf59602 Human cel
126	33	82.5	466	9	ACH33036	Ach33036 Human end	c 199	33	82.5	2611	4	AAI59032	Aai59032 Human pol
127	33	82.5	474	2	AAV66763	Aav66763 Pathogen	c 200	33	82.5	2611	5	ADQ99255	Adq99255 DNA encod
128	33	82.5	501	12	ADI34835	Adi34835 Bovine pr	c 201	33	82.5	2611	9	ADB49015	Adb49015 Novel hum
129	33	82.5	526	13	ACN61641	Acn61641 Cotton gy	c 202	33	82.5	2679	6	ABL61970	Abli61970 Colon ade
130	33	82.5	565	13	ACN60146	Acn60146 Cotton gy	c 203	33	82.5	2679	6	ABN97278	Abn97278 Gene #377
131	33	82.5	571	13	ACN62686	Acn62686 Cotton de	c 204	33	82.5	2680	13	ADX55069	Adx55069 Plant ful
132	33	82.5	575	12	ADJ67315	Adj67315 Human ova	c 205	33	82.5	2840	4	AAI60818	Aai60818 Human pol
133	33	82.5	575	12	ADP81034	Adp81034 Human ova	c 206	33	82.5	2982	13	ADK14549	Adk14549 Plant ful
134	33	82.5	617	6	ABQ57009	Abq57009 Human col	c 207	33	82.5	3161	12	ADQ63982	Adq63982 Novel hum
135	33	82.5	628	3	AAK50850	Aak50850 Arabidops	c 208	33	82.5	3240	4	AAH54379	Aah54379 S. epider
136	33	82.5	633	3	AAK40534	Aak40534 Arabidops	c 209	33	82.5	3600	2	AAK25180	Aak25180 HIV-1 gro
137	33	82.5	656	13	ACN62428	Acn62428 Cotton de	c 210	33	82.5	3802	10	ABT16587	Abt16587 Ethylene
138	33	82.5	718	11	ACN82227	Acn82227 Breast ca	c 211	33	82.5	4183	4	ABL15626	Abli15626 Drosophil
139	33	82.5	771	10	ACP67051	AcP67051 Phototrab	c 212	33	82.5	4600	4	ABL09067	Abli09067 Drosophil
140	33	82.5	786	9	ADA29868	Ada29868 DNA encod	c 213	33	82.5	4654	4	ABL03141	Abli03141 Drosophil
141	33	82.5	819	6	ABK73662	Abk73662 Bacillus	c 214	33	82.5	5871	4	AAF26075	Aaf26075 FIV gp140
142	33	82.5	825	6	ABN66859	Abn66859 Streptoco	c 215	33	82.5	6259	6	AAD29720	Aad29720 Plaemid p
143	33	82.5	870	8	ABZ51090	Abz51090 Aspergill	c 216	33	82.5	6682	2	AAK12993	Aak12993 Enterococ
144	33	82.5	880	11	ACN82602	Acn82602 Breast ca	c 217	33	82.5	6682	6	ABG98788	Abg98788 Enterococ
145	33	82.5	1095	2	AAV74888	Aav74888 Staphyloc	c 218	33	82.5	8863	4	ABL09066	Abli09066 Drosophil
146	33	82.5	1104	10	ACF71505	AcF71505 Phototrab	c 219	33	82.5	9642	4	ABL03140	Abli03140 Drosophil
147	33	82.5	1107	8	ACA27902	Aca27902 Prokaryot	c 220	33	82.5	9656	4	AAK46523	Aak46523 Tumour su
148	33	82.5	1143	5	ADL63180	Adl63180 Human ova	c 221	33	82.5	9656	6	ABN80212	Abn80212 Tumour che
149	33	82.5	1150	8	ADA73250	Ada73250 Rice gene	c 222	33	82.5	9829	3	AAZ35271	Aaz35271 Soybean r
150	33	82.5	1150	11	ACI35631	AcI35631 Rice stre	c 223	33	82.5	12578	4	AAK46660	Aak46660 Tumour su

516	31	77.5	292	10	ADD28551	Add28551 Becherich	c 589	31	77.5	590	12	ACH67609	Ach67609 Human gen
517	31	77.5	292	10	ADD28549	Add28549 Becherich	c 590	31	77.5	591	13	ADT05037	Adt05037 Haenophil
518	31	77.5	292	10	ADD28548	Add28548 Becherich	c 591	31	77.5	591	13	ADT05037	Adt05037 Haenophil
519	31	77.5	300	9	ADB093355	Adb093355 Human bra	c 592	31	77.5	602	14	ADY66533	Ady66533 S. manson
520	31	77.5	326	2	AAQ60206	Aaq60206 Human bra	c 593	31	77.5	608	14	ACL60817	ACL60817 Human col
521	31	77.5	330	6	ABN24329	Abn24329 Human ORF	c 594	31	77.5	609	10	ADC76392	Adc76392 DNA homol
522	31	77.5	335	4	AAK87681	Aak87681 Human imm	c 595	31	77.5	609	10	ADC75748	Adc75748 DNA homol
523	31	77.5	355	4	AAK87679	Aak87679 Human imm	c 596	31	77.5	609	10	ADKS8339	Adk58339 Plant DNA
524	31	77.5	369	14	ACL70421	ACL70421 M. xanthu	c 597	31	77.5	620	10	ADKS4273	Adk54273 Plant DNA
525	31	77.5	382	4	AAI13195	Aai13195 Probe #31	c 598	31	77.5	620	10	AAI93444	Aai93444 Human pol
526	31	77.5	382	4	ABA54895	Abas54895 Human fce	c 599	31	77.5	625	11	ADL65876	Adl65876 C. glucan
527	31	77.5	382	4	AAI134547	Aai134547 Probe #32	c 600	31	77.5	633	8	ACA28882	ACA28882 Prokaryot
528	31	77.5	382	4	ABA44448	Abas44448 Human bre	c 601	31	77.5	643	6	ABN62507	Abn62507 Human can
529	31	77.5	382	4	ABA24661	Abas24661 Probe #31	c 602	31	77.5	645	6	ACC00599	Acc00599 A. thalia
530	31	77.5	382	4	AAK28621	Aak28621 Human bon	c 603	31	77.5	657	6	ABQ38275	Abq38275 Oligonucl
531	31	77.5	382	4	AAK03171	Aak03171 Human bra	c 604	31	77.5	657	6	ABQ38274	Abq38274 Oligonucl
532	31	77.5	382	5	AAI03100	Aai03100 Probe #30	c 605	31	77.5	665	8	ACA57501	ACA57501 Human adi
533	31	77.5	382	6	ABSO3132	Abso3132 Human gen	c 606	31	77.5	671	10	ACF67292	Acf67292 Photorhab
534	31	77.5	384	5	AAK64674	Aak64674 Novel hum	c 607	31	77.5	681	2	AAT80823	Aat80823 Staphyloc
535	31	77.5	394	6	ABL65544	AbL65544 Lung canc	c 608	31	77.5	688	13	ADS11400	Adsl11400 Human the
536	31	77.5	397	8	ABZ18864	Abz18864 Group III	c 609	31	77.5	694	13	ADX30754	Adx30754 Plant ful
537	31	77.5	399	13	ADR61184	Adr61184 Cotton cd	c 610	31	77.5	698	3	AAF13335	Aaf13335 Aspergill
538	31	77.5	416	5	ABV18065	Abv18065 Human pro	c 611	31	77.5	698	13	ADU57376	Adu57376 Aspergill
539	31	77.5	420	6	ABL80854	AbL80854 Human ova	c 612	31	77.5	700	14	ADZ95379	Adz95379 Aspergill
540	31	77.5	421	8	ABX52124	Abx52124 Bovine ES	c 613	31	77.5	705	6	AAH92099	Aah92099 Human inf
541	31	77.5	424	13	ACR88238	Acr88238 Human SIR	c 614	31	77.5	715	6	ABK75977	Abk75977 Bacillus
542	31	77.5	432	13	ADS51159	Ads51159 Bacterial	c 615	31	77.5	717	14	ADV43070	Adv43070 Human psy
543	31	77.5	438	12	ADP93393	Adp93393 Cotton ex	c 616	31	77.5	727	4	AAH16844	Aah16844 Human cDN
544	31	77.5	442	13	ACF91126	Acf91126 Human SIR	c 617	31	77.5	727	6	ABQ37497	Abq37497 Oligonucl
545	31	77.5	450	5	ABV52477	Abv52477 Human pro	c 618	31	77.5	727	6	ABQ37496	Abq37496 Oligonucl
546	31	77.5	457	9	ACH39935	Ach39935 Human fce	c 619	31	77.5	740	4	AAH06966	Aah06966 Human cDN
547	31	77.5	457	12	ACH83766	Ach83766 Human gen	c 620	31	77.5	741	2	AAV24764	Aav24764 H. pylori
548	31	77.5	461	6	ABL81143	AbL81143 Human ova	c 621	31	77.5	744	6	ABQ47118	Abq47118 Oligonucl
549	31	77.5	464	4	AAI15476	Aai15476 Probe #54	c 622	31	77.5	744	6	ABQ47119	Abq47119 Oligonucl
550	31	77.5	464	4	ABAS7320	Abas7320 Human fce	c 623	31	77.5	760	4	AAH06488	Aah06488 Human cDN
551	31	77.5	464	4	AAI136864	Aai136864 Probe #55	c 624	31	77.5	761	2	AAV44292	Aav44292 Human sec
552	31	77.5	464	4	ABA26870	Abas26870 Probe #53	c 625	31	77.5	761	5	AAF98466	Aaf98466 Human cDN
553	31	77.5	464	4	AAK03951	Aak03951 Human bon	c 626	31	77.5	766	13	ADX35593	Adx35593 Plant ful
554	31	77.5	464	4	AAK05356	Aak05356 Human bra	c 627	31	77.5	771	6	ABQ20108	Abq20108 Oligonucl
555	31	77.5	464	4	ABSO30629	Abso30629 Human liv	c 628	31	77.5	771	6	ABQ20109	Abq20109 Oligonucl
556	31	77.5	464	6	ABSO5699	Abso5699 Human gen	c 629	31	77.5	783	10	ADC92389	Adc92389 E. faeciu
557	31	77.5	469	2	AAV89203	Aav89203 EST Clone	c 630	31	77.5	793	13	ADX33141	Adx33141 Plant ful
558	31	77.5	471	4	AAH29040	Aah29040 Drosophil	c 631	31	77.5	800	10	ACF67031	Acf67031 Photorhab
559	31	77.5	472	2	AAT19048	Aat19048 Human gen	c 632	31	77.5	805	2	ADR01696	Adr01696 A. gossyp
560	31	77.5	474	6	AAK33682	Aak33682 Human eos	c 633	31	77.5	805	13	ADX59336	Adx59336 Plant ful
561	31	77.5	492	11	ADJ12090	Adj12090 Maize cDN	c 634	31	77.5	837	13	ADT17604	Adt17604 Rice cDN
562	31	77.5	498	13	ADQ56175	Adq56175 Novel can	c 635	31	77.5	842	4	ABK42285	Abk42285 Genomic s
563	31	77.5	500	9	ACH43843	Ach43843 Human fce	c 636	31	77.5	842	9	ADB60441	Adb60441 Connectiv
564	31	77.5	500	10	ADC03538	Adc03538 Human Na/	c 637	31	77.5	849	10	ABZ40310	Abz40310 N. gonorr
565	31	77.5	512	6	ABQ33745	Abq33745 Oligonucl	c 638	31	77.5	852	3	AAZ53292	Aaz53292 Neisseria
566	31	77.5	512	6	ABQ33744	Abq33744 Oligonucl	c 639	31	77.5	852	3	AAZ53291	Aaz53291 Neisseria
567	31	77.5	512	8	ABZ73128	Abz73128 Rice leaf	c 640	31	77.5	852	3	AAZ53290	Aaz53290 Neisseria
568	31	77.5	513	13	ADR62533	Adr62533 Cotton cd	c 641	31	77.5	852	3	AAZ53290	Aaz53290 Neisseria
569	31	77.5	516	5	AAH66212	Aah66212 C. glutami	c 642	31	77.5	867	10	ADH85155	Adh85155 Enterococ
570	31	77.5	516	8	ACA00421	Aca00421 C. glutam	c 643	31	77.5	876	3	AAZ53292	Aaz53292 Neisseria
571	31	77.5	518	5	AAH88529	Aah88529 DNA encod	c 644	31	77.5	899	13	ADX5416	Adx5416 Plant ful
572	31	77.5	518	13	ACF81257	Acf81257 Human SIR	c 645	31	77.5	921	8	ACA27485	ACA27485 Prokaryot
573	31	77.5	531	6	ABN93364	Abn93364 Staphyloc	c 646	31	77.5	942	10	ADH82200	Adh82200 Enterococ
574	31	77.5	531	13	ADS04391	Ads04391 Staphyloc	c 647	31	77.5	963	8	ACA27036	ACA27036 Prokaryot
575	31	77.5	533	5	ABV51764	Abv51764 Human pro	c 648	31	77.5	982	2	AAZ42241	Aaz42241 Human nor
576	31	77.5	537	9	ABO7587	AbO7587 Alloiococ	c 649	31	77.5	993	14	ABE66969	AbE66969 Rice geno
577	31	77.5	540	6	ABQ91277	Abq91277 M. capsul	c 650	31	77.5	996	11	ABD09371	Abd09371 Pseudomon
578	31	77.5	540	10	ADQ56380	Adq56380 Toxicity-	c 651	31	77.5	1000	14	ABE85724	AbE85724 Human pho
579	31	77.5	546	12	ACH70066	Ach70066 Human gen	c 652	31	77.5	1001	3	AAZ57814	Aaz57814 Arachidon
580	31	77.5	550	6	ABV96679	Abv96679 Human pan	c 653	31	77.5	1020	6	ABN70631	Abn70631 Streptoco
581	31	77.5	554	6	ABN66643	Abn66643 Streptoco	c 654	31	77.5	1023	8	ACA50665	ACA50665 Prokaryot
582	31	77.5	567	13	ADU69347	Adu69347 S agalact	c 655	31	77.5	1028	14	ABE66816	AbE66816 Rice geno
583	31	77.5	567	13	ADV84580	Adv84580 Streptoco	c 656	31	77.5	1032	5	AAZ58510	Aaz58510 DNA encod
584	31	77.5	568	12	ACH71601	Ach71601 Human gen	c 657	31	77.5	1035	13	ADT17734	Adt17734 Plant cDN
585	31	77.5	573	13	ADQ54949	Adq54949 Novel can	c 658	31	77.5	1039	13	ADX50032	Adx50032 Plant ful
586	31	77.5	579	3	AAZ29496	Aaz29496 Incyte cl	c 659	31	77.5	1047	6	ABN67010	Abn67010 Streptoco
587	31	77.5	582	4	AAI01896	Aai01896 Human rep	c 660	31	77.5	1104	4	ABL21647	AbL21647 Drosophil
588	31	77.5	582	4	ABL97189	AbL97189 Human tes	c 661	31	77.5	1112	13	ADX33358	Adx33358 Plant ful
										1116	13	ADO83346	Ado83346 Plant ful

662	31	77.5	1133	8	ACA333372	Aca333372 Prokaryot	c 735	31	77.5	1440	2	AAV24965	Aav24965 H. pylori
663	31	77.5	1136	3	AAC48066	Aac48066 Zee may's	c 736	31	77.5	1440	2	AAV57599	Aav57599 H. pylori
664	31	77.5	1137	13	ADX50811	Adx50811 Plant ful	c 737	31	77.5	1446	4	AAS51764	Aas51764 Staphyloc
665	31	77.5	1139	13	ADO82156	Ado82156 Plant ful	c 738	31	77.5	1452	4	AAS05172	Aas05172 Tactu's
666	31	77.5	1141	13	ADO81670	Ado81670 Plant ful	c 739	31	77.5	1453	3	AAZ96744	Aaz96744 Nuclear t
667	31	77.5	1143	5	AAH66534	Aah66534 C. glutami	c 740	31	77.5	1453	13	ADX13256	Adx13256 Plant ful
668	31	77.5	1148	13	ADO82720	Ado82720 Plant ful	c 741	31	77.5	1461	2	AAV24911	Aav24911 H. pylori
669	31	77.5	1149	13	ADX31611	Adx31611 Plant ful	c 742	31	77.5	1461	2	AAV58000	Aav58000 H. pylori
670	31	77.5	1150	13	ADX59788	Adx59788 Plant ful	c 743	31	77.5	1461	8	ACA47362	Aca47362 Prokaryot
671	31	77.5	1150	13	ADX64413	Adx64413 Plant ful	c 744	31	77.5	1464	8	ACF72787	Acf72787 Staphyloc
672	31	77.5	1153	13	ADX46270	Adx46270 Plant ful	c 745	31	77.5	1467	4	AAS54829	Aas54829 Staphyloc
673	31	77.5	1167	6	ABN59693	Abn59693 Novel hum	c 746	31	77.5	1467	8	ACA20120	Aca20120 Prokaryot
674	31	77.5	1167	13	ADX33417	Adx33417 Plant ful	c 747	31	77.5	1487	4	AAS36398	Aas36398 Human car
675	31	77.5	1170	3	AAC50356	Aac50356 Arabidops	c 748	31	77.5	1487	4	AAS36399	Aas36399 Human car
676	31	77.5	1171	13	ADX59841	Adx59841 Plant ful	c 749	31	77.5	1487	4	AAS36397	Aas36397 Human car
677	31	77.5	1174	3	ACA38057	Aac38057 Arabidops	c 750	31	77.5	1487	10	ADBA7093	Ade47093 Human car
678	31	77.5	1179	8	ACA33243	Aca33243 Prokaryot	c 751	31	77.5	1487	10	ADBA7091	Ade47091 Human car
679	31	77.5	1182	4	ABA82969	Aba82969 Enterococ	c 752	31	77.5	1487	10	ADBA7092	Ade47092 Human car
680	31	77.5	1198	13	ADX32640	Adx32640 Plant ful	c 753	31	77.5	1487	13	ADJ08510	Adj08510 Human car
681	31	77.5	1207	3	AAZ29481	Aaz29481 DNA encod	c 754	31	77.5	1487	13	ADJ08509	Adj08509 Human car
682	31	77.5	1210	10	ADB52426	Adb52426 Primary r	c 755	31	77.5	1487	13	ADJ08511	Adj08511 Human car
683	31	77.5	1218	13	ADX30853	Adx30853 Plant ful	c 756	31	77.5	1497	10	ADC91756	Adc91756 E. faeciu
684	31	77.5	1230	6	AAV39750	Aad39750 Human A2a	c 757	31	77.5	1551	8	ACA24907	Aca24907 Prokaryot
685	31	77.5	1230	6	AAV39751	Aad39751 Human A2a	c 758	31	77.5	1568	13	ADX52470	Aca52470 Plant ful
686	31	77.5	1233	13	ADX53184	Adx53184 Plant ful	c 759	31	77.5	1617	2	AAV63624	Aav63624 CDNA enco
687	31	77.5	1238	12	ADQ76826	Adq76826 Human w1	c 760	31	77.5	1617	2	AAV63624	Aav63624 CDNA enco
688	31	77.5	1238	12	AAQ48415	Aaq48415 Human A2a	c 761	31	77.5	1617	3	AAA09430	Aaa09430 M. alpina
689	31	77.5	1239	2	AAT07649	Aat07649 Human ade	c 762	31	77.5	1617	12	ADP89958	Adp89958 M. alpina
690	31	77.5	1239	2	AAT29930	Aat29930 Human ven	c 763	31	77.5	1617	13	ADR20160	Adr20160 Mortierel
691	31	77.5	1239	2	AAT00645	Aat00645 Human A2a	c 764	31	77.5	1642	3	ACF74292	Acf74292 Human sec
692	31	77.5	1239	6	AAV39749	Aad39749 Human A2a	c 765	31	77.5	1645	12	ADK14351	Adk14351 Yarrowia
693	31	77.5	1239	6	AAV39748	Aad39748 Human A2a	c 766	31	77.5	1645	12	ADK14351	Adk14351 Yarrowia
694	31	77.5	1239	6	AAV25371	Aad25371 Human ade	c 767	31	77.5	1645	12	ADK14351	Adk14351 Yarrowia
695	31	77.5	1239	12	ADQ29798	Ado29798 Human GPC	c 768	31	77.5	1657	13	ADX55059	Adx55059 Plant ful
696	31	77.5	1239	12	ADQ29798	Ado29798 Human GPC	c 769	31	77.5	1658	4	AAV26013	Aav26013 Human CDN
697	31	77.5	1242	11	ABD09471	Abd09471 Pseudomon	c 770	31	77.5	1658	8	ABX73354	Abx73354 Human nov
698	31	77.5	1245	4	AAV52296	Aas52296 E. coli D	c 771	31	77.5	1680	13	ADO82689	Ado82689 Plant ful
699	31	77.5	1245	5	AAH81461	Aah81461 Escherich	c 772	31	77.5	1686	12	ADQ76862	Adq76862 Adenosine
700	31	77.5	1245	2	ACA20422	Aca20422 Human sec	c 773	31	77.5	1710	8	ACF74326	Acf74326 Staphyloc
701	31	77.5	1257	2	AAV20422	Aad20422 Human sec	c 774	31	77.5	1716	6	ABI99953	Abi99953 Human DPB
702	31	77.5	1257	10	ADG90026	Adg90026 Novel hum	c 775	31	77.5	1716	12	ADI43009	Adi43009 Plant tra
703	31	77.5	1257	10	ADG90025	Adg90025 Human CDN	c 776	31	77.5	1716	12	ADO03140	Ado03140 Soybean o
704	31	77.5	1257	14	ADY25365	Ady25365 Novel hum	c 777	31	77.5	1719	14	ADW00767	Adw00767 Nucleotid
705	31	77.5	1260	6	ABL91303	Ab191303 Chlamydia	c 778	31	77.5	1722	13	ADS46035	Ads46035 Bacterial
706	31	77.5	1296	6	ABN69255	Abn69255 Streptoco	c 779	31	77.5	1735	13	ADO81987	Ado81987 Plant ful
707	31	77.5	1320	4	ABL24385	Ab124385 Streptoco	c 780	31	77.5	1737	12	ADQ76866	Adq76866 Adenosine
708	31	77.5	1329	3	AAC44055	Aac44055 Zee may's	c 781	31	77.5	1737	12	ADQ76860	Adq76860 Adenosine
709	31	77.5	1335	8	ACAS4365	Aca54365 Prokaryot	c 782	31	77.5	1767	4	AAH14697	Aah14697 Human CDN
710	31	77.5	1338	6	ABN70729	Abn70729 Streptoco	c 783	31	77.5	1768	4	ABL05441	Ab105441 Drosophil
711	31	77.5	1341	4	AAV07047	Aas07047 DNA encod	c 784	31	77.5	1779	8	ACA25210	Aca25210 Prokaryot
712	31	77.5	1341	13	ADV84144	Adv84144 Streptoco	c 785	31	77.5	1788	12	ADQ76864	Adq76864 Adenosine
713	31	77.5	1353	8	ACA42315	Aca42315 Prokaryot	c 786	31	77.5	1794	14	ADW16458	Adw16458 Eucalyptu
714	31	77.5	1358	2	AAV74651	Aav74651 Staphyloc	c 787	31	77.5	1798	4	AAS32724	Aas32724 Human gen
715	31	77.5	1366	4	AAS59774	Aas59774 Propionib	c 788	31	77.5	1821	10	ABZ25668	Abz25668 Human KIA
716	31	77.5	1366	8	ACF64703	Acf64703 Propionib	c 789	31	77.5	1825	5	AAS45198	Aas45198 CDNA enco
717	31	77.5	1371	14	AEBA7303	Aeb47303 Fatty aci	c 790	31	77.5	1829	6	ABL34408	Ab134408 Human imm
718	31	77.5	1374	3	AAZ47129	Aaz47129 Fungal de	c 791	31	77.5	1840	13	ADX14457	Adx14457 Plant ful
719	31	77.5	1374	5	AAF25234	Aaf25234 Nucleotid	c 792	31	77.5	1873	4	AAH13643	Aah13643 Human CDN
720	31	77.5	1374	14	ADU77738	Adu77738 M. alpina	c 793	31	77.5	1882	6	ABL90479	Ab190479 Human pol
721	31	77.5	1374	14	ADU82421	Adu82421 M. alpina	c 794	31	77.5	1883	13	ADX54013	Adx54013 Plant ful
722	31	77.5	1374	14	ADU81422	Adu81422 Mortierel	c 795	31	77.5	1906	6	AAD25395	Aad25395 Human ade
723	31	77.5	1376	2	AAV74329	Aav74329 Staphyloc	c 796	31	77.5	1906	6	AAD25370	Aad25370 Human ade
724	31	77.5	1383	13	ADO83738	Ado83738 Plant ful	c 797	31	77.5	1944	12	ADN72230	Adn72230 Thale cre
725	31	77.5	1394	6	ABV77910	Abv77910 Hypoxia-i	c 798	31	77.5	1983	13	ADX52990	Adx52990 Plant ful
726	31	77.5	1395	11	ABD10339	Abd10339 Pseudomon	c 799	31	77.5	2000	6	ABZ16042	Abz16042 Arabidops
727	31	77.5	1407	6	ABL90165	Ab190165 Human pol	c 800	31	77.5	2000	6	ABZ15614	Abz15614 Arabidops
728	31	77.5	1423	13	ADX15340	Adx15340 Plant ful	c 801	31	77.5	2000	8	ADA72546	Ada72546 Rice gene
729	31	77.5	1431	6	ABN70025	Abn70025 Streptoco	c 802	31	77.5	2000	10	ACC60851	Acc60851 Gene sequ
730	31	77.5	1434	4	AAS06369	Aas06369 Streptoco	c 803	31	77.5	2000	10	ADK62209	Adk62209 Disease t
731	31	77.5	1434	11	ABD09570	Abd09570 Pseudomon	c 804	31	77.5	2000	11	ACL35403	Ac135403 Rice stre
732	31	77.5	1437	11	ABD10536	Abd10536 Pseudomon	c 805	31	77.5	2073	5	AAS45010	Aas45010 CDNA enco
733	31	77.5	1439	3	AAC56201	Aac56201 Pinus rad	c 806	31	77.5	2088	4	AAH54099	Aah54099 S. epider
734	31	77.5	1440	2	AAV24964	Aav24964 H. pylori	c 807	31	77.5	2091	10	ADC90882	Adc90882 E. faeciu

c 808	31	77.5	2108	3	AAC83326	Aac83326 Human PAR	c 881	31	77.5	2791	14	ADY52842	Ady52842 Mus muscu
c 809	31	77.5	2108	6	ABK92160	Abk92160 Prostate	882	31	77.5	2796	13	ADP24487	Adp24487 PRO polyyp
c 810	31	77.5	2112	11	ADM01622	Adm01622 Human CDN	883	31	77.5	2811	5	AAS63169	Aas63169 Human pur
c 811	31	77.5	2160	6	ABL54576	AbL54576 Mouse per	884	31	77.5	2859	9	ACA98937	AcA98937 cDNA enco
c 812	31	77.5	2167	9	ADA19505	Ada19505 Dog plas	885	31	77.5	2875	8	ABQ83871	AbQ83871 Human MDD
c 813	31	77.5	2167	14	ADZ70051	Adz70051 Canine p1	886	31	77.5	2903	4	AAK51891	AaK51891 Human pol
c 814	31	77.5	2186	6	ABL34639	AbL34639 Human met	887	31	77.5	2917	4	ABL21648	AbL21648 Drosophil
c 815	31	77.5	2186	6	ABL70616	AbL70616 Chemical	888	31	77.5	2932	4	ABL29634	AbL29634 Drosophil
c 816	31	77.5	2186	7	ADS9900	AdS9900 Complemen	889	31	77.5	2959	8	ABQ83881	AbQ83881 Human MDD
c 817	31	77.5	2191	2	AAQ87949	AaQ87949 Canine p1	c 890	31	77.5	2988	3	AAA35135	AaA35135 Human ade
c 818	31	77.5	2191	2	AAT80581	AaT80581 Canine p1	c 891	31	77.5	2988	3	AAF21257	AaF21257 Human low
c 819	31	77.5	2191	2	AAT66138	AaT66138 Canine PA	c 892	31	77.5	2988	6	ABZ35318	AbZ35318 Human gen
c 820	31	77.5	2191	2	AAT87065	AaT87065 Canine p1	c 893	31	77.5	2988	10	ABA296951	AbA296951 Human nuc
c 821	31	77.5	2191	2	AAT96129	AaT96129 Canine p1	c 894	31	77.5	2988	10	ACA56759	AcA56759 Human sig
c 822	31	77.5	2191	2	AAX08477	AaX08477 Canine p1	c 895	31	77.5	2988	11	ABD20800	AbD20800 Human pul
c 823	31	77.5	2191	2	AAV08551	AaV08551 Canine PA	c 896	31	77.5	2988	12	ADI56555	AdI56555 Human pol
c 824	31	77.5	2191	3	AAA10878	AaA10878 Canine PA	c 897	31	77.5	3007	4	AAS32779	AaS32779 Human gen
c 825	31	77.5	2191	3	AAZ24254	AaZ24254 Canine PA	c 898	31	77.5	3013	12	ADQ63655	AdQ63655 Novel hum
c 826	31	77.5	2191	3	AAAS9593	AaA9593 cDNA enco	c 899	31	77.5	3017	4	AAH14566	AaH14566 Human CDN
c 827	31	77.5	2191	4	AAD04157	AaD04157 Dog plas	900	31	77.5	3017	14	ADX06966	AdX06966 Cyclin-de
c 828	31	77.5	2191	4	AAC89071	AaC89071 Platelet-	901	31	77.5	3017	14	ADZ09668	AdZ09668 Human bre
c 829	31	77.5	2191	5	AAD24733	AaD24733 Canine p1	c 902	31	77.5	3038	4	ABL16776	AbL16776 Drosophil
c 830	31	77.5	2266	4	ABL09225	AbL09225 Drosophil	903	31	77.5	3042	12	ADI24466	AdI24466 Human mod
c 831	31	77.5	2283	10	ADB95023	AdB95023 A. thalia	904	31	77.5	3042	12	ADP21330	AdP21330 Gene Arp1
c 832	31	77.5	2313	13	ADW47573	Adt47573 Bacterial	c 905	31	77.5	3072	4	ABL27349	AbL27349 Drosophil
c 833	31	77.5	2351	14	ADW10445	Adw10445 Colon pro	c 906	31	77.5	3085	4	ABL16797	AbL16797 Drosophil
c 834	31	77.5	2357	12	ADQ23082	AdQ23082 Human sof	c 907	31	77.5	3111	2	AAQ87260	AaQ87260 Flea sodi
c 835	31	77.5	2361	12	ADQ76836	AdQ76836 A2A adeno	908	31	77.5	3132	8	ACA29420	AcA29420 Prokaryot
c 836	31	77.5	2377	12	ADQ23475	AdQ23475 Human sof	909	31	77.5	3153	8	ACA46504	AcA46504 Prokaryot
c 837	31	77.5	2383	3	AAAS35134	AaA35134 Human ade	910	31	77.5	3156	6	ABN92705	AbN92705 Staphyloc
c 838	31	77.5	2383	3	AAE21256	AaE21256 Human low	911	31	77.5	3156	13	ADS03435	AdS03435 Staphyloc
c 839	31	77.5	2383	6	ABZ35598	AbZ35598 Human gen	c 912	31	77.5	3198	12	AAV01883	AaV01883 Human nel
c 840	31	77.5	2383	10	ABZ96950	AbZ96950 Human nuc	c 913	31	77.5	3198	6	ABL64482	AbL64482 Lung can
c 841	31	77.5	2383	11	ABD20799	AbD20799 Human pul	c 914	31	77.5	3198	6	ABK64420	AbK64420 Human ben
c 842	31	77.5	2403	8	ABZ42616	AbZ42616 Human ade	c 915	31	77.5	3198	8	ACC69974	AcC69974 Human nel
c 843	31	77.5	2403	14	ADV15459	Adv15459 DNA enco	c 916	31	77.5	3198	12	ADQ18200	AdQ18200 Human sof
c 844	31	77.5	2403	14	ADZ48789	AdZ48789 Insulin s	c 917	31	77.5	3198	13	ADQ76518	AdQ76518 Nucleotid
c 845	31	77.5	2420	4	ABL24386	AbL24386 Drosophil	c 918	31	77.5	3198	13	ADR25390	AdR25390 Breast ca
c 846	31	77.5	2448	2	AAV01882	AaV01882 Human nel	c 919	31	77.5	3198	13	ADP56111	AdP56111 Human PRO
c 847	31	77.5	2448	8	ACC69973	AcC69973 Human nel	c 920	31	77.5	3198	13	ADV34943	AdV34943 Human hom
c 848	31	77.5	2448	13	ADQ76517	AdQ76517 Nucleotid	c 921	31	77.5	3198	14	ADY19905	AdY19905 DNA enco
c 849	31	77.5	2451	13	ADR32028	AdR32028 Human NEL	c 922	31	77.5	3198	14	ADY15077	AdY15077 DNA enco
c 850	31	77.5	2451	14	ADV43996	Adv43996 Human psy	c 923	31	77.5	3211	4	AAH54847	AaH54847 S. epider
c 851	31	77.5	2484	6	ABN68372	AbN68372 Streptoco	c 924	31	77.5	3213	8	ACA54251	AcA54251 Prokaryot
c 852	31	77.5	2487	8	ACA50513	AcA50513 Prokaryot	c 925	31	77.5	3234	6	ABA01698	AbA01698 Methanosa
c 853	31	77.5	2487	8	ABX63444	AbX63444 Human CDN	c 926	31	77.5	3277	4	ABL08584	AbL08584 Drosophil
c 854	31	77.5	2501	8	ABZ10219	AbZ10219 Haematopo	927	31	77.5	3287	2	AAQ55140	AaQ55140 Staphyloc
c 855	31	77.5	2501	14	ADV98053	Adv98053 Bisulfite	928	31	77.5	3300	2	AAQ62589	AaQ62589 Thermotog
c 856	31	77.5	2551	5	ABV29879	Abv29879 Human pro	929	31	77.5	3348	8	ABX63664	AbX63664 Human CDN
c 857	31	77.5	2551	5	ABV22623	Abv22623 Human pro	c 930	31	77.5	3438	2	AAT95247	AaT95247 M. catar
c 858	31	77.5	2551	5	ABV23996	Abv23996 Human pro	931	31	77.5	3465	13	ADU87322	AdU87322 T. mariti
c 859	31	77.5	2551	5	ABV29856	Abv29856 Human pro	932	31	77.5	3526	4	AAI58305	AaI58305 Human pol
c 860	31	77.5	2551	5	ABV23971	Abv23971 Human pro	933	31	77.5	3526	5	ADQ98512	AdQ98512 DNA enco
c 861	31	77.5	2551	5	ABV28445	Abv28445 Human pro	934	31	77.5	3526	9	ADB48272	AdB48272 Novel hum
c 862	31	77.5	2551	5	ABV25375	Abv25375 Human pro	c 935	31	77.5	3532	11	ADM02135	AdM02135 Human CDN
c 863	31	77.5	2551	5	ABV25214	Abv25214 Human pro	936	31	77.5	3556	4	RAK52875	RaK52875 Human pol
c 864	31	77.5	2562	13	ADT15866	Adt15866 Plant cDN	937	31	77.5	3556	13	AAI60091	AaI60091 Human pol
c 865	31	77.5	2567	5	AAQ56001	AaQ56001 Angiotens	938	31	77.5	3556	13	ACN40597	AcN40597 Tumour-as
c 866	31	77.5	2575	2	AAQ54641	AaQ54641 Human A2a	939	31	77.5	3557	10	ADB85135	AdB85135 Rat NA,K-K-
c 867	31	77.5	2575	2	AAV73321	AaV73321 Human A2a	940	31	77.5	3557	10	ADD68431	AdD68431 Chicken s
c 868	31	77.5	2575	9	ADA19420	Ada19420 Adenosine	c 941	31	77.5	3563	4	RAH54008	RaH54008 S. epider
c 869	31	77.5	2592	13	ADT05243	Adt05243 Haemophil	c 942	31	77.5	3582	8	ABX49976	AbX49976 Human CDN
c 870	31	77.5	2636	10	AAI52245	AaI52245 cDNA enco	c 943	31	77.5	3588	4	ABL06015	AbL06015 Drosophil
c 871	31	77.5	2639	2	AAV52934	AaV52934 Pig trans	944	31	77.5	3717	4	ABL21646	AbL21646 Drosophil
c 872	31	77.5	2648	4	AAAS9683	AaA9683 Propionib	945	31	77.5	3727	12	ADJ87624	AdJ87624 Nervous s
c 873	31	77.5	2648	8	ACF64612	AcF64612 Propionib	c 946	31	77.5	3751	14	ACL64123	AcL64123 M. xanthu
c 874	31	77.5	2669	2	AAQ56925	AaQ56925 Pig TGF-b	947	31	77.5	3768	4	ABL05440	AbL05440 Drosophil
c 875	31	77.5	2671	2	AAQ03303	AaQ03303 Entire po	c 948	31	77.5	3784	4	ABL24384	AbL24384 Drosophil
c 876	31	77.5	2676	2	AAQ02819	AaQ02819 cDNA sequ	949	31	77.5	3791	10	ADB62393	AdB62393 Human CDN
c 877	31	77.5	2690	4	AAH10005	AaH10005 Drosophil	c 950	31	77.5	3824	4	RAC84430	RaC84430 Human tes
c 878	31	77.5	2749	4	AAH14476	AaH14476 Human CDN	c 951	31	77.5	3824	12	ADK60466	AdK60466 Angiogene
c 879	31	77.5	2766	4	AAAS01571	AaA01571 Human sec	c 952	31	77.5	3824	12	ADP73089	AdP73089 Angiogene
c 880	31	77.5	2791	2	AAK60597	AaK60597 Murine OX	c 953	31	77.5	3824	12	ADP73089	AdP73089 Angiogene

c 954 31 77.5 3887 6 ABQ99287
 c 955 31 77.5 3902 12 ADQ63047
 c 956 31 77.5 3916 8 ABX49964
 c 957 31 77.5 3924 5 ABX63188
 c 958 31 77.5 4120 5 AAS90070
 c 959 31 77.5 4135 6 ABK89314
 c 960 31 77.5 4135 12 ADJ57736
 c 961 31 77.5 4161 4 AAS4897
 c 962 31 77.5 4213 6 ABQ70936
 c 963 31 77.5 4233 8 ABZ10136
 c 964 31 77.5 4233 10 ABZ10240
 c 965 31 77.5 4233 10 ADE84150
 c 966 31 77.5 4234 2 AAX13250
 c 967 31 77.5 4234 6 AAS99045
 c 968 31 77.5 4237 3 AAC76647
 c 969 31 77.5 4293 12 ADN05365
 c 970 31 77.5 4353 10 ADG42133
 c 971 31 77.5 4353 10 ADF74252
 c 972 31 77.5 4366 13 ADU82759
 c 973 31 77.5 4369 4 AAH18383
 c 974 31 77.5 4369 13 ADQ83948
 c 975 31 77.5 4571 8 ABE57531
 c 976 31 77.5 4572 14 ADS58525
 c 977 31 77.5 4682 4 ABL09224
 c 978 31 77.5 4695 10 ACF71169
 c 979 31 77.5 4755 14 AEA39467
 c 980 31 77.5 4755 14 AEA39470
 c 981 31 77.5 4755 14 AEA39468
 c 982 31 77.5 4755 14 AEA39441
 c 983 31 77.5 4755 14 AEA39469
 c 984 31 77.5 4755 14 AEB28639
 c 985 31 77.5 4755 14 AEB28640
 c 986 31 77.5 4755 14 AEB28638
 c 987 31 77.5 4755 14 AEB28637
 c 988 31 77.5 4755 14 AEB28611
 c 989 31 77.5 4817 12 ADQ59191
 c 990 31 77.5 4828 4 AAH18238
 c 991 31 77.5 4828 12 ADP10546
 c 992 31 77.5 4988 4 AAS41973
 c 993 31 77.5 5109 10 ACC48816
 c 994 31 77.5 5109 13 ADV40868
 c 995 31 77.5 5176 8 ABX34580
 c 996 31 77.5 5184 6 ABK31484
 c 997 31 77.5 5184 6 ABL70453
 c 998 31 77.5 5184 6 AAS61420
 c 999 31 77.5 5207 10 ADC30042
 1000 31 77.5 5231 10 ADD19033

ALIGNMENTS

RESULT 1
 ID ACD97670 standard; cDNA; 108 BP.

XX ACD97670;
 XX
 XX 23-SEP-2003 (first entry)
 DT Human colon cancer cell expressed cDNA #6082.

DE Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KW agriculture; food crop genome; resistance gene; retrovirus;
 KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.

OS Homo sapiens.
 XX
 XX US2002155438-A1.
 PN
 XX
 XX 24-OCT-2002.

XX 27-SEP-1999; 99US-00406117.
 XX 20-NOV-1998; 98US-00196716.
 PR (SIMP/) SIMPSON A J G.
 XX (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 XX Simpson AJG, Neto ED, Brentani RR;
 XX WPI; 2003-182626/18.
 XX
 XX Determining open reading frames of genome of an organism e.g. a human
 XX suffering from cancer involves use of single oligonucleotide primer at
 XX low stringency for preparing single-stranded cDNA from mRNA of
 XX individual.
 XX
 XX Example 9; Page 866; 959pp; English.
 XX
 CC The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (1) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (1), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. Further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer. To determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a fetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present. Defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 XX
 SQ Sequence 108 BP; 18 A; 33 C; 22 G; 35 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.04 Length: 108
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-10 (1-9) x ACD97670 (1-108)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 81 TTCTGGGTATGTTTAGCCCTGATA 107

RESULT 2

ADU11677

ID ADU11677 standard; DNA; 475 BP.

XX AC ADU11677;
XX DT 27-JAN-2005 (first entry)
XX DE Solid tumour prognosis gene seqid 2116.
XX KW cytostatic; gene therapy; expression profile; solid tumour;
XX KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX OS Unidentified.
XX PN WO2004097052-A2.
XX PD 11-NOV-2004.
XX PF 29-APR-2004; 2004WO-US013587.
XX PR 29-APR-2003; 2003US-0466067P.
XX PR 23-JAN-2004; 2004US-0538246P.
XX PA (AMHP) WYETH.
XX PA (STRA/) STRAHS A.
XX PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
XX PI Immerman F, Dornier AJ;
XX DR WPI; 2004-804779/79.
XX PT A method, useful for prognosing and treating solid tumor, comprises
XX PT comparing an expression profile of a gene expressed in peripheral blood
XX PT mononuclear cells to a reference expression profile of a gene.
XX PS Disclosure; Page; 111pp; English.
XX CC The invention describes a method comprising comparing an expression
XX CC profile of at least one gene in a peripheral blood sample of a patient to
XX CC at least one reference expression profile of the at least one gene, where
XX CC the patient has a solid tumour, and each of the gene is differentially
XX CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
XX CC of patients as compared to PBMCs of a second class of patients, where
XX CC both the first and second classes of patients have the solid tumour, and
XX CC each of the first and second classes is a subcluster formed by an
XX CC unsupervised clustering analysis of gene expression profiles in PBMCs of
XX CC a population of patients who have the solid tumour, and where the
XX CC majority of the first class of patients has a first clinical outcome, and
XX CC the majority of the second class of patients has a second clinical
XX CC outcome. Also described are: a system comprising (i) a memory or a
XX CC storage medium including data that represent an expression profile of at
XX CC least one gene in a peripheral blood sample of a patient who has a solid
XX CC tumour, (ii) at least another storage medium including data that
XX CC represent at least one reference expression profile of the gene, (iii) a
XX CC program capable of comparing the expression profile to the reference
XX CC expression profile, and (iv) a processor capable of executing the
XX CC program, where expression levels of the gene in peripheral blood
XX CC mononuclear cells of patients who have the solid tumour correlate with
XX CC clinical outcomes of the patients; and a nucleic acid or protein array
XX CC comprising concentrated probes for solid tumour prognosis genes, where
XX CC each of the solid tumour prognosis genes is differentially expressed in
XX CC PBMCs of a first class of patients as compared to PBMCs of a second class
XX CC of patients, where both the first and second classes of patients have a
XX CC solid tumour, and where the first class of patients has a first clinical
XX CC outcome, and the second class of patients has a second clinical outcome.
XX CC The method, system, and array are useful for prognosing and treating
XX CC solid tumours. This sequence represents a solid tumour prognosis gene of
XX CC the invention. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;
Alignment Scores: 22.5 Length: 475
Pred. No.:
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0
US-10-774-176-10 (1-9) x ADU11677 (1-475)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 351 TTCCTGGGTATGTTTAGCCCTGATA 377
RESULT 3
ABT07721
ID ABT07721 standard; DNA; 927 BP.
XX AC ABT07721;
XX DT 14-NOV-2002 (first entry)
XX DE Breast cancer-associated gene sequence 29.
XX KW Gene; ds; breast cancer; breast cancer-associated gene sequence;
XX KW drug development; pharmacogenetics; biosensor development.
XX OS Unidentified.
XX PN WO200259377-A2.
XX PD 01-AUG-2002.
XX PF 24-JAN-2002; 2002WO-US002242.
XX PR 24-JAN-2001; 2001US-0263965P.
XX PR 02-FEB-2001; 2001US-0265928P.
XX PR 09-APR-2001; 2001US-00829472.
XX PR 09-APR-2001; 2001US-0282698P.
XX PR 04-MAY-2001; 2001US-0288590P.
XX PR 29-MAY-2001; 2001US-0294443P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Mack DH, Gish KC, Afar D;
XX WPI; 2002-583738/62.
XX DR N-PSDB; ABJ05564.
XX PT Detecting a breast cancer-associated transcript in a patient's cell,
XX PT useful for diagnosing breast cancer, comprises contacting a biological
XX PT sample with a polynucleotide that selectively hybridizes with breast
XX PT cancer nucleic acids.
XX PS Claim 9; Page 372; 414pp; English.
XX CC The invention comprises a method of detecting a breast cancer-associated
XX CC transcript in a cell from a patient. The method of the invention involves
XX CC contacting a biological sample from the patient with a nucleotide that
XX CC hybridises to one of the 69 breast cancer-associated gene sequences shown
XX CC in the specification. The method of the invention is useful in the
XX CC diagnosis or prognosis of breast cancer, and for detecting genes that are
XX CC up or down-regulated in breast cancer cells. Genes identified by the
XX CC method of the invention can be used in diagnostic purposes and also as
XX CC targets for screening for therapeutic compounds that modulate breast
XX CC cancer (e.g. hormones or antibodies). Identification of genes that are
XX CC over or under expressed in breast cancer can additionally provide high-
XX CC resolution, high-sensitivity datasets which can be used in the areas of
XX CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
XX CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
XX CC represent the 69 breast cancer-associated gene sequences of the invention
XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
Alignment Scores:

Pred. No.: 48.9 Length: 927
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x ABT07721 (1-927)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 730 TTCTGGGTATTGTTTGGCCCTGATA 756

RESULT 4

ABX76333

ID ABX76333 standard; DNA; 927 BP.

XX AC ABX76333;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polynucleotide #197.

XX KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
XX KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.

XX PR 18-APR-2001; 2001US-0284770P.

XX PR 10-MAY-2001; 2001US-0290492P.

XX PR 09-NOV-2001; 2001US-0339245P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 29-NOV-2001; 2001US-0334370P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX DR WPI; 2003-093161/08.

XX DR P-PSDB; ABUS6604.

XX PT Detecting a lung cancer-associated transcript in a cell from a patient
XX PT for treating lung cancer, by contacting a biological sample from the
XX PT patient with a polynucleotide that exhibits increased or decreased
XX PT expression in lung cancer.

XX PS Claim 22; Page 336; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated
XX CC transcript in a cell from a patient, comprising contacting a biological
XX CC sample from the patient with a polynucleotide that selectively hybridizes
XX CC to a sequence that is at least 80 % identical to a gene that exhibits
XX CC increased or decreased expression in lung cancer samples. Lung cancer-
XX CC associated polynucleotides and polypeptides are used for identifying a
XX CC compound that modulates a lung cancer-associated polypeptide, for
XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung
XX CC cancer in a patient and for treating a mammal having lung cancer by
XX CC administering a modulatory compound identified. The methods are useful
XX CC for treating lung cancer, such as small cell lung cancer, non-small cell
XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful

CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 48.9 Length: 927
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x ABX76333 (1-927)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 730 TTCTGGGTATTGTTTGGCCCTGATA 756

RESULT 5

ADB80503

ID ADB80503 standard; DNA; 927 BP.

XX AC ADB80503;

XX DT 04-DEC-2003 (first entry)

XX DE Ovarian cancer-associated transcript #34.

XX KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
XX KW post-operative chemotherapy; radiation therapy; tumour prognosis;
XX KW pre-cancerous lesion detection; ds; gene.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..927

XX FT /*tag= a

XX PN WO2002102235-A2.

XX PD 27-DEC-2002.

XX PF 18-JUN-2002; 2002WO-US019297.

XX PR 18-JUN-2001; 2001US-0299234P.

XX PR 27-AUG-2001; 2001US-0315287P.

XX PR 05-SEP-2001; 2001US-0317544P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Mack DH, Gish KC;

XX DR WPI; 2003-167431/16.

XX DR P-PSDB; ADB80504.

XX PT Detecting an ovarian cancer-associated transcript in a cell from a
XX PT patient, comprises contacting a biological sample from the patient with a
XX PT polynucleotide that hybridizes to an ovarian cancer gene.

XX PS Claim 10; Page 297; 332pp; English.

XX CC The invention relates to a method of detecting an ovarian cancer-
XX CC associated transcript in a cell from a patient, by contacting a
XX CC biological sample from the patient with a polynucleotide that selectively
XX CC hybridizes to a sequence at least 80% identical to any of one of 80
XX CC nucleic acid sequences given in the specification. The method is useful
XX CC in diagnosing ovarian cancer and in identifying and using agents and/or
XX CC targets that inhibit ovarian cancer. The nucleic acid molecule,

CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selection mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.

XX
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores: 48.9 Length: 927
 Pred. No.: 40.00 Matches: 9
 Score: 100.0% Conservative: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 100.0% Gaps: 0
 DB: 10

US-10-774-176-10 (1-9) x ADB80503 (1-927)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 |||||
 Db 730 TTCTGGGTATGTTTAGCCCTGATA 756

RESULT 6

ADN38723
 ID ADN38723 standard; cDNA; 927 BP.

XX
 AC ADN38723;

DT 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

DE Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

OS WO2003042661-A2.

PN 22-MAY-2003.

PD 13-NOV-2002; 2002WO-US036810.

PF 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.

PR 10-JAN-2002; 2002US-0347349P.

PR 08-FEB-2002; 2002US-0355250P.

PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-036809P.

PR 04-APR-2002; 2002US-0370110P.

PR 12-APR-2002; 2002US-037246P.

PR 05-JUN-2002; 2002US-0386614P.

PR 16-JUL-2002; 2002US-0396839P.

XX
 DR
 DR
 XX

WPI: 2003-468649/44.
 P-PSDB; ADN38724.

Determining the presence or absence of a pathological cell in a patient,
 useful for diagnosing, prognosing or treating cancer, comprises detecting
 a nucleic acid in a biological sample.

XX
 XX Claim 8; SEQ ID NO 41; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 whose expression is upregulated or downregulated in specific cancers or
 other diseases such as angiogenic or fibrotic disorders, and to methods
 of determining the presence or absence of a pathological cell in a
 patient by detecting a nucleic acid at least 80% identical to those of
 the invention or by detecting a polypeptide of the invention. The
 invention also relates to expression vectors and host cells comprising a
 nucleic acid of the invention; antibodies which specifically bind a
 polypeptide of the invention; use of such antibodies for drug targeting;
 and methods of screening for modulators of activity or expression of the
 polypeptides and nucleic acids. The nucleic acids, polypeptides,
 antibodies and methods are useful for diagnosing, prognosing and treating
 cancer and other conditions such as psoriasis, ischaemia, heart disease,
 atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 neovascularisation syndromes, scarring and uterine fibroids. They may
 also be useful in wound healing and in contraception. The present
 sequence represents a nucleic acid sequence of the invention.

XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 48.9 Length: 927
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

US-10-774-176-10 (1-9) x ADN38723 (1-927)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 |||||
 Db 730 TTCTGGGTATGTTTAGCCCTGATA 756

RESULT 7

AAD56198
 ID AAD56198 standard; DNA; 973 BP.

XX
 AC AAD56198;

XX
 DT 07-AUG-2003 (first entry)

XX Human LRRCAPS related DNA #5.

XX Human; p53 pathway; Leucine rich repeat capricious related protein;
 LRRCAPS; cancer; gene therapy; ds.

XX Homo sapiens.

XX WO2003035831-A2.

XX 01-MAY-2003.

XX 21-OCT-2002; 2002WO-US033540.

XX 22-OCT-2001; 2001US-0338733P.

XX 15-FEB-2002; 2002US-0357600P.

XX 01-MAR-2002; 2002US-0361196P.

XX (EXEL-) EXELIXIS INC.

XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 Francis-Lang H, Friedman L;

XX WPI; 2003-421410/39.
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX Example 5; Page 74-75; 99pp; English.
 XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 51.7 Length: 973
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x AAD56198 (1-973)
 QY 1 PheLeuGlyTleValLeuAlaLeuIle 9
 Db 745 TTCTGGGTATTGTTTTCAGCCCTGATA 771

RESULT 8
 ID ABV99349
 XX ABV99349 standard; DNA; 1156 BP.
 AC ABV99349;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Human NOV8a coding sequence.
 XX
 KW Human; anti-HIV; cytostatic; antidiabetic; antiaesthatic; cachexia; AIDS;
 KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
 KW neutropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene-
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200272771-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US0007288.
 XX
 PR 08-MAR-2001; 2001US-0274101P.
 PR 08-MAR-2001; 2001US-0274194P.
 PR 08-MAR-2001; 2001US-0274381P.
 PR 08-MAR-2001; 2001US-0274322P.
 PR 12-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275578P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR

13-MAR-2001; 2001US-0275601P.
 14-MAR-2001; 2001US-0276000P.
 16-MAR-2001; 2001US-0276776P.
 19-MAR-2001; 2001US-0276994P.
 20-MAR-2001; 2001US-0277239P.
 20-MAR-2001; 2001US-0277321P.
 20-MAR-2001; 2001US-0277327P.
 20-MAR-2001; 2001US-0277338P.
 21-MAR-2001; 2001US-0277791P.
 22-MAR-2001; 2001US-0277833P.
 23-MAR-2001; 2001US-0278152P.
 26-MAR-2001; 2001US-0278894P.
 27-MAR-2001; 2001US-0278999P.
 28-MAR-2001; 2001US-0279036P.
 30-MAR-2001; 2001US-0279344P.
 30-MAR-2001; 2001US-0279995P.
 30-MAR-2001; 2001US-0280233P.
 02-APR-2001; 2001US-0280802P.
 02-APR-2001; 2001US-0280822P.
 02-APR-2001; 2001US-0280900P.
 04-APR-2001; 2001US-0281194P.
 13-APR-2001; 2001US-0283675P.
 30-APR-2001; 2001US-0287424P.
 02-MAY-2001; 2001US-0288066P.
 03-MAY-2001; 2001US-0288342P.
 15-MAY-2001; 2001US-0288528P.
 16-MAY-2001; 2001US-0291190P.
 16-MAY-2001; 2001US-0291099P.
 16-MAY-2001; 2001US-0291240P.
 30-MAY-2001; 2001US-0294485P.
 31-MAY-2001; 2001US-0294889P.
 31-MAY-2001; 2001US-0294899P.
 18-JUN-2001; 2001US-0299027P.
 19-JUN-2001; 2001US-0299303P.
 19-JUN-2001; 2001US-0299310P.
 10-JUL-2001; 2001US-0304354P.
 31-JUL-2001; 2001US-0309198P.
 16-AUG-2001; 2001US-0312903P.
 10-SEP-2001; 2001US-0318462P.
 12-SEP-2001; 2001US-0318770P.
 27-SEP-2001; 2001US-0325430P.
 27-SEP-2001; 2001US-0325681P.
 18-OCT-2001; 2001US-0330380P.
 31-OCT-2001; 2001US-0335301P.
 14-NOV-2001; 2001US-0332172P.
 14-NOV-2001; 2001US-0332271P.
 14-NOV-2001; 2001US-0332272P.
 14-NOV-2001; 2001US-0333184P.
 21-NOV-2001; 2001US-0333272P.
 03-DEC-2001; 2001US-0332094P.
 03-DEC-2001; 2001US-0337426P.
 04-DEC-2001; 2001US-0338092P.
 04-DEC-2001; 2001US-0337185P.
 03-JAN-2002; 2002US-0345705P.
 08-MAR-2002; 2002US-00093463.
 (CURA-) CURAGEN CORP.
 Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 Zhong M;
 WPI; 2002-732824/79.
 P-PSDB; ABP70071.
 New NOVX polypeptides and polynucleotides, useful for preventing,
 diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 disorders, and asthma.

PS Claim 16; Page 114-115; 61pp; English.

XX The present invention relates to new isolated proteins (NOVX) and their

CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is

CC any number from 1 to 48. The NOVX proteins and coding sequences are

CC useful in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, preferably a NOVX-associated disorder.

CC The NOVX coding sequences and proteins are useful for treating,

CC preventing or diagnosing diseases such as metabolic disorders, diabetes,

CC obesity, infectious diseases, anorexia, cancer-associated cachexia,

CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's

CC disease, immune disorders, haematopoietic disorders, cardiovascular

CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic

CC disturbances associated with obesity, metabolic syndrome X or wasting

CC disorders associated with chronic diseases or various cancers. The NOVX

CC coding sequences and proteins may also be used as targets for the

CC identification of small molecules that modulate or inhibit e.g.

CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,

CC wound healing and angiogenesis, in gene therapy, in generation of

CC antibodies that bind immunospecifically to NOVX substances for use in

CC therapeutic or diagnostic methods

XX

SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	63.1	Length:	1156
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-10 (1-9) x ABV99349 (1-1156)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9

DB 961 TTCTGGGTATTGTTTAGCCCTGATA 987

RESULT 9

ID ABK87175

XX ABK87175 standard; cDNA; 1260 BP.

AC ABK87175;

XX

XX 07-OCT-2002 (first entry)

XX

XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.

XX

XX Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;

XX cell proliferative disorder; infection; inflammatory condition;

XX cancer immunotherapy; foetal cell; maternal blood; cytostatic;

XX foetal abnormality; foetal sex determination; gene; ss.

XX

OS Felis sp.

XX

XX Key Location/Qualifiers

XX CDS 1..1260

XX

XX /*tag= a

XX /product= "5T4 protein"

XX

XX WO200238612-A2.

XX

XX 16-MAY-2002.

XX

XX 13-NOV-2001; 2001WO-GB005004.

XX

XX 13-NOV-2000; 2000WO-GB004317.

XX

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX

XX Myers K, Drury N, Carroll M;

XX

XX WPI; 2002-557449/59.

XX

DR P-PSDB; AAU98694.

XX

PT Novel canine or feline 5T4 polypeptide and polynucleotides encoding the

PT polypeptide, useful in preparation of vaccine for treating and/or

PT preventing cancer in a subject, preferably a dog or cat.

XX

PS Claim 4; Page 68; 68pp; English.

XX

CC The present invention relates to the isolation of canine and feline

CC oncofoetal leucine-rich glycoproteins known as 5T4, and the

CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in

CC a significant proportion of tumours. The sequences of the invention are

CC useful in a pharmaceutical composition for the prevention and/or

CC treatment of tumours or other diseases associated with cell

CC proliferation, infections, and inflammatory conditions in animals,

CC preferably dogs or cats. The compositions may also be used for cancer

CC immunotherapy in these animals. The sequences of the invention may also

CC be used in diagnostic kits for rapid, reliable, sensitive, and specific

CC measurement and localisation of 5T4 in extracts of plasma, urine,

CC tissues, and in cell culture media. Antibodies specific for the 5T4

CC protein are useful for isolating foetal cells from maternal blood. The

CC isolation process may form part of a diagnostic method e.g. the foetal

CC cells may then be subject to biochemical or genetic sampling used for

CC testing foetal abnormalities, or to determine the sex of the foetus(es).

CC The present sequence encodes feline 5T4 protein

XX

SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	69.8	Length:	1260
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-10 (1-9) x ABK87175 (1-1260)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9

DB 1069 TTCTAGGTATTGTTTAGCCCTGATA 1095

RESULT 10

ID ADB97513

XX ADB97513 standard; DNA; 1260 BP.

XX

XX ADB97513;

XX

XX 04-DEC-2003 (first entry)

XX

XX Feline 5T4 antigen DNA.

XX

XX Major Histocompatibility Complex class I peptide epitope; MHC;

XX 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;

XX cytostatic; cancer; feline; gene; ds.

XX

XX Unidentified.

XX

XX Key Location/Qualifiers

XX CDS 1..1260

XX

XX /*tag= a

XX /product= "Feline 5T4 antigen protein"

XX

XX WO2003068816-A1.

XX

XX 21-AUG-2003.

XX

XX 13-FEB-2003; 2003WO-GB000670.

XX

XX 13-FEB-2002; 2002GB-00003419.

XX

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX

PI Carroll M, Kingsman S, Redchenko I;
 XX WPI; 2003-637141/60.
 DR P-PSDB; ADB97520.
 XX
 XX New major histocompatibility complex class I peptide epitopes from human
 PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.
 XX
 XX Disclosure; Page 67; 73pp; English.
 XX
 CC The invention relates to a novel Major Histocompatibility Complex (MHC)
 CC class I peptide epitope of the 5T4 antigen. The invention further
 CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
 CC acid to a cell; a cell pulsed with the 5T4 epitope; a polypeptide of the
 CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 5T4 epitope and/or its encoding nucleic acid; a method
 CC comprising detecting the presence of the 5T4 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 5T4 epitope in conjunction with an MHC class
 CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC particularly cancer. The detection method is useful for diagnosing or
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline
 CC 5T4 antigen coding DNA of the invention.
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 69.8 Length: 1260
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-10 (1-9) x ADB97513 (1-1260)
 QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 DB 1069 TTCTAGGTATTGTTTAGCCCTGATA 1095
 RESULT 11
 ADB97452
 ID ADB97452 standard; DNA; 1260 BP.
 XX
 AC ADB97452;
 XX
 XX 04-DEC-2003 (first entry)
 XX
 DE DNA encoding feline 5T4 protein.
 KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
 KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 CD 1..1260
 FT /tag= a
 FT /product= "Feline 5T4 antigen protein"
 FT
 XX WO2003068815-A2.
 XX
 PD 21-AUG-2003.

XX 13-FEB-2003; 2003WO-GB000618.
 XX
 XX 13-FEB-2002; 2002GB-00003420.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Carroll M, Harrop R, Kingsman S;
 XX
 XX WPI; 2003-663795/62.
 DR P-PSDB; ADB97455.
 XX
 XX New Major Histocompatibility Complex class II peptide epitope of 5T4,
 PT useful for manufacturing a medicament for diagnosing, preventing and/or
 PT treating a disease, e.g. cancer.
 XX
 XX Disclosure; Page 49; 63pp; English.
 XX
 CC The invention relates to a Major Histocompatibility Complex (MHC) class
 CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
 CC clone has a cytostatic activity, as it is useful in manufacturing a
 CC medicament for preventing and/or treating a disease, particularly cancer.
 CC The methods are useful for detecting T-cells capable of specifically
 CC recognising a peptide epitope in conjunction with an MHC molecule, for
 CC diagnosing or monitoring the progression of a cancerous disease, or for
 CC detecting the presence of a peptide or nucleic acid using an agent. The
 CC MHC class II peptide epitope of the invention can be used in gene therapy
 CC or as part of a vaccine. This polynucleotide sequence represents the DNA
 CC coding for the feline 5T4 protein.
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 69.8 Length: 1260
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-10 (1-9) x ADB97452 (1-1260)
 QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 DB 1069 TTCTAGGTATTGTTTAGCCCTGATA 1095
 RESULT 12
 AAA27058
 ID AAA27058 standard; DNA; 1263 BP.
 XX
 AC AAA27058;
 XX
 XX 22-AUG-2000 (first entry)
 XX
 DE Human 5T4 tumour-associated antigen gene.
 XX
 KW Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
 KW ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200029428-A2.
 XX
 PD 25-MAY-2000.
 XX
 XX 18-NOV-1999; 99WO-GB003859.
 XX
 XX 18-NOV-1998; 98GB-00025303.
 PR 27-JAN-1999; 99GB-00001739.
 PR 30-JUL-1999; 99GB-00017995.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.

```

XX Carroll MW, Myers KA;
PI WPI; 2000-387735/33.
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
PT response useful in vaccinating against and in treating tumors.
XX Example 2; Page 78; 79pp; English.
XX The present sequence encodes the human 5T4 tumour-associated antigen
CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
CC carcinomas but has a highly restricted expression pattern in normal adult
CC tissues. It appears to be strongly correlated to metastasis in colorectal
CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
CC induced were inoculated with a virus expression vector containing the
CC present sequence. The 5T4 antigen was shown to be effective at eliciting
CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
CC the antigen and the antigen itself can be used to elicit an immune
CC response, preferably CTL or an antibody response in a subject
XX
SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 69.9 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-10 (1-9) x AAF89736 (1-1263)
Qy 1 PheLeuGlyIleValLeuAlaLeuile 9
Db 1072 TTCTGGGTATTGTTTTCAGCCCTGATA 1098

RESULT 13
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX AAF89736;
XX 23-JUL-2001 (first entry)
XX Nucleotide sequence of canine 5T4 protein.
XX Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
KW hypersensitivity; autoimmune disease; central nervous system disorder;
KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
KW Helicobacter-related disease; immune disorder; ss.
XX
OS Canis sp.
FH Key Location/Qualifiers
FT CDS 1..1263
FT /*tag= a
FT /product= "5T4"
XX
PN WO200136486-A2.
XX 25-MAY-2001.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX 18-NOV-1999; 99WO-GB003859.
XX 15-FEB-2000; 2000GB-00003527.
XX 02-MAR-2000; 2000GB-00005071.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX

PI Kingman A, Kingman SM, Bebbington CR, Carroll MW, Ellard FM;
XX Myers KA;
DR WPI; 2001-343805/36.
DR P-PSDB; AAB83839.
XX Use of single chain antibody capable of recognizing a disease associated
PT molecule for manufacturing a medicament for preventing and/or treating a
XX disease condition associated with disease associated molecule.
XX Disclosure; Fig 26; 118pp; English.
XX The specification describes the use of a single chain antibody (ScFv),
CC which is capable of recognizing a disease associated molecule in the
CC manufacture of a medicament for the prevention and treatment of a disease
CC condition. The ScFv antibody is useful in the manufacture of a
CC medicament, for affecting a disease in vivo, for preparing a
CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
CC treatment of a disease. The ScFv antibody is also useful for treating
CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
CC diseases, cancers, central nervous system disorders including Parkinson's
CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
CC related diseases, and other immune disorders. The present sequence
CC encodes a 5T4 protein, which is used to produce ScFv of the invention
XX
SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 69.9 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-10 (1-9) x AAF89736 (1-1263)
Qy 1 PheLeuGlyIleValLeuAlaLeuile 9
Db 1072 TTCTAGGTATTGTTTTCAGCCCTGATA 1098

RESULT 14
ABK87174
ID ABK87174 standard; cDNA; 1263 BP.
XX ABK87174;
XX 07-OCT-2002 (first entry)
XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
XX Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX
OS Canis sp.
FH Key Location/Qualifiers
FT CDS 1..1263
FT /*tag= a
FT /product= "5T4 protein"
XX
PN WO200238612-A2.
XX 16-MAY-2002.
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX

```

```

XX Myers K, Drury N, Carroll M;
PI WPI; 2002-557449/59.
XX P-PSDB; AAU98693.
DR
XX Novel canine or feline 574 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 1; Page 67; 68pp; English.
P8
XX The present invention relates to the isolation of canine and feline
CC oncofetal leucine-rich glycoproteins known as 574, and the
CC polynucleotide sequences encoding them. The 574 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 574 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 574
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus(es).
CC The present sequence encodes canine 574 protein
XX
SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 69.9 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x ABK87174 (1-1263)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1072 TTCCTAGGTATGCTTAGCCCTGATA 1098

RESULT 15
AA27059
ID AAA27059 standard; DNA; 1281 BP.
XX
XX AAA27059;
AC
XX
XX 22-AUG-2000 (first entry)
DT
XX
DE Mouse 574 tumour-associated antigen gene.
XX
KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW da.
XX
XX Mus musculus.
OS
XX
XX WO200029428-A2.
PN
XX
XX 25-MAY-2000.
PD
XX
XX 18-NOV-1999; 99WO-GB003859.
PF
XX
XX 18-NOV-1998; 98GB-00025303.
PR
XX 27-JAN-1999; 99GB-00001739.
PR
XX 30-JUL-1999; 99GB-00017995.
PR
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA

XX Carroll MW, Myers KA;
PI WPI; 2000-387735/33.
XX
XX Tumor associated antigen, 574 capable of eliciting cytotoxic T-lymphocyte
PT response useful in vaccinating against and in treating tumors.
XX
XX Example 2; Page 78; 79pp; English.
P8
XX The present sequence encodes the mouse 574 tumour-associated antigen
CC (TAA). The TAA 574 is a glycoprotein which is widely expressed in
CC carcinomas but has a highly restricted expression pattern in normal adult
CC tissues. It appears to be strongly correlated to metastasis in colorectal
CC and gastric cancer. 574 antigen may therefore be useful in tumour
CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
CC induced were inoculated with a virus expression vector containing the
CC present sequence. The 574 antigen was shown to be effective at eliciting
CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
CC the antigen and the antigen itself can be used to elicit an immune
CC response, preferably CTL or an antibody response in a subject. The
CC present sequence appears in GenBank at accession number AJ012160
XX
SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 71.1 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-10 (1-9) x AAA27059 (1-1281)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1090 TTCCTAGGTATGCTTAGCCCTGATA 1116

RESULT 16
AAD56199
ID AAD56199 standard; DNA; 1331 BP.
XX
XX AAD56199;
AC
XX
XX 07-AUG-2003 (first entry)
DT
XX
DE Human LRRCAPS related DNA #6.
XX
KW Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
XX
XX Homo sapiens.
OS
XX
XX WO2003035831-A2.
PN
XX
XX 01-MAY-2003.
PD
XX
XX 21-OCT-2002; 2002WO-US033540.
PF
XX
XX 22-OCT-2001; 2001US-0338733P.
PR
XX 15-FEB-2002; 2002US-0357600P.
PR
XX 01-MAR-2002; 2002US-0361196P.
PR
XX (EXEL-) EXELIXIS INC.
PA
XX
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
XX WPI; 2003-421410/39.
DR
XX
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich

```

PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX Disclosure; Page 75-76; 99pp; English.
PS
CC The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 74.3 Length: 1331
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x AAD56199 (1-1331)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1102 TTCTGGGTATTGTTTAGCCCTGATA 1128
|||||

RESULT 17

ADJ56299
ID ADJ56299 standard; cDNA; 2020 BP.
XX
AC ADJ56299;
DT 06-MAY-2004 (first entry)
XX
DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.
XX human; differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX
OS Homo sapiens.
XX
PN US2003119009-A1.
XX
PD 26-JUN-2003.
XX
PF 25-FEB-2002; 2002US-00084817.
XX
PR 23-FEB-2001; 2001US-0270784P.
XX
(STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLOW/) PLOW S E.
PA (SHOH/) SHOHEIT J M.
XX
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX
DR WPI; 2003-635698/60.
XX
XX New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
PS Claim 1; SEQ ID NO 105; 27pp; English.
XX
XX This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to

CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX

SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 121 Length: 2020
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-10 (1-9) x ADJ56299 (1-2020)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1142 TTCTGGGTATTGTTTAGCCCTGATA 1168
|||||

RESULT 18

ACC51052
ID ACC51052 standard; cDNA; 2053 BP.
XX
AC ACC51052;
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated cDNA sequence SEQ ID NO:192.
XX Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
XX
PN WO2003003906-A2.
XX
PD 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US021338.
XX
PR 03-JUL-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Aziz N;
XX
DR WPI; 2003-201532/19.
XX
DR P-PSDB; ABR48236.
XX
XX Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
PS Claim 6; Page 296; 307pp; English.
XX
XX The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises

CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridizes to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 123 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x ACC51052 (1-2053)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
 |||||
 Db 1156 TTCTGGGTATTGTTTGGCCCTGATA 1182

RESULT 19

ABX76332
 ID ABX76332 standard; DNA; 2053 BP.

XX AC ABX76332;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polynucleotide #196.

XX KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.

XX PR 18-APR-2001; 2001US-0284770P.

XX PR 10-MAY-2001; 2001US-0290492P.

XX PR 09-NOV-2001; 2001US-0339245P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 29-NOV-2001; 2001US-0334370P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (BOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX XX WPI; 2003-093161/08.

XX DR P-PSDB; ABUS6603.

XX PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

PS Claim 22; Page 335; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 123 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x ABX76332 (1-2053)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 1156 TTCTGGGTATTGTTTGGCCCTGATA 1182

RESULT 20

AAD56197

ID AAD56197 standard; DNA; 2053 BP.

XX AC AAD56197;

XX DT 07-AUG-2003 (first entry)

XX DE Human LRRCAPS DNA #11.

XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO2003035831-A2.

XX PD 01-MAY-2003.

XX PF 21-OCT-2002; 2002WO-US033540.

XX PR 22-OCT-2001; 2001US-0338733P.

XX PR 15-FEB-2002; 2002US-0357600P.

XX PR 01-MAR-2002; 2002US-0361196P.

XX PA (EXEL-) EXELIXIS INC.

XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;

XX DR WPI; 2003-421410/39.

XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.

XX Example 5; Page 73-74; 99pp; English.
 XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores: 123 Length: 2053
 Pred. No.: 40.00 Matches: 9
 Score: 100.0% Conservative: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 100.0% Gaps: 0
 DB: 8

US-10-774-176-10 (1-9) x AAD56197 (1-2053)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
 |||||
 Db 1156 TTCTGGGTATTGTTTGTAGCCCTGATA 1182

RESULT 21

AAD56200
 ID AAD56200 standard; DNA; 2053 BP.

XX AAD56200;
 XX 07-AUG-2003 (first entry)
 DT Human LRRCAPS DNA #12.
 DE Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX Homo sapiens.

OS WO2003035831-A2.
 XX 01-MAY-2003.

PD 21-OCT-2002; 2002WO-US033540.
 XX 22-OCT-2001; 2001US-0338733P.
 XX 15-FEB-2002; 2002US-0357600P.
 XX 01-MAR-2002; 2002US-0361196P.

XX (EXEL-) EXELIXIS INC.

XX Belvin M, Schleithoff L, Plozman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX WPI; 2003-421410/39.
 DR Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX Disclosure; Page 76-77; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the

CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores: 123 Length: 2053
 Pred. No.: 40.00 Matches: 9
 Score: 100.0% Conservative: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 100.0% Gaps: 0
 DB: 8

US-10-774-176-10 (1-9) x AAD56200 (1-2053)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
 |||||
 Db 1156 TTCTGGGTATTGTTTGTAGCCCTGATA 1182

RESULT 22

ADN38721
 ID ADN38721 standard; cDNA; 2053 BP.

XX ADN38721;
 XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.

KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.

XX Homo sapiens.
 XX WO2003042661-A2.
 XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.
 XX 13-NOV-2001; 2001US-0350666P.
 XX 21-NOV-2001; 2001US-0332464P.
 XX 29-NOV-2001; 2001US-0334393P.
 XX 03-DEC-2001; 2001US-0335394P.
 XX 14-DEC-2001; 2001US-0340376P.
 XX 08-JAN-2002; 2002US-0347211P.
 XX 10-JAN-2002; 2002US-0347349P.
 XX 08-FEB-2002; 2002US-0355250P.
 XX 13-FEB-2002; 2002US-0356714P.
 XX 20-FEB-2002; 2002US-0359077P.
 XX 29-MAR-2002; 2002US-0368809P.
 XX 04-APR-2002; 2002US-0370110P.
 XX 12-APR-2002; 2002US-0372246P.
 XX 05-JUN-2002; 2002US-0386614P.
 XX 16-JUL-2002; 2002US-0396839P.
 XX 22-JUL-2002; 2002US-039775P.
 XX 22-JUL-2002; 2002US-0397845P.
 XX 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX WPI; 2003-468649/44.

DR P-PSDB; ADN38722.

XX Determining the presence or absence of a pathological cell in a patient,

PT useful for diagnosing, prognosing or treating cancer, comprises detecting

PT a nucleic acid in a biological sample.

XX Claim 9; SEQ ID NO 39; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)

CC whose expression is upregulated or downregulated in specific cancers or

CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a

CC patient by detecting a nucleic acid at least 80% identical to those of

CC the invention or by detecting a polypeptide of the invention. The

CC invention also relates to expression vectors and host cells comprising a

CC nucleic acid of the invention; antibodies which specifically bind a

CC polypeptide of the invention; use of such antibodies for drug targeting;

CC and methods of screening for modulators of activity or expression of the

CC polypeptides and nucleic acids. The nucleic acids, polypeptides,

CC antibodies and methods are useful for diagnosing, prognosing and treating

CC cancer and other conditions such as psoriasis, ischaemia, heart disease,

CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal

CC neovascularisation syndromes, scarring and uterine fibroids. They may

CC also be useful in wound healing and in contraception. The present

CC sequence represents a nucleic acid sequence of the invention.

XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	123	Length:	2053
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-774-176-10 (1-9) x ADN38721 (1-2053)

QY 1 PheLeuGlyTleValLeuAlaLeuile 9

DB 1156 TTCCTGGGTATTGTTTAGCCCTGATA 1182

RESULT 23

ADL06473

ID ADL06473 standard; cDNA; 2053 BP.

XX ADL06473;

XX 20-MAY-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #53.

XX Human; tumour-associated antigenic target; TAT; cell death; tumour;

XX cancer; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004016225-A2.

XX 26-FEB-2004.

XX 19-AUG-2003; 2003WO-US025892.

XX 19-AUG-2002; 2002US-0404809P.

XX 21-AUG-2002; 2002US-0405645P.

XX 23-SEP-2002; 2002US-0413192P.

XX 15-OCT-2002; 2002US-0419008P.

XX 15-NOV-2002; 2002US-0426847P.

XX 02-JUL-2003; 2003US-0484959P.

XX (GETH) GENENTECH INC.

XX Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;

PI Spencer SD, Wu TD, Zhang Z;

XX WPI; 2004-257144/24.

XX P-PSDB; ADL06552.

XX New antibody that binds to a tumor-associated antigenic target (TAT)

PT polypeptide, useful for preparing a composition for diagnosing or

PT treating cancer.

XX Claim 1; SEQ ID NO 53; 319pp; English.

XX The present invention relates to the isolation of human tumour-associated

CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also

CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is

CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a

CC humanised antibody. It is conjugated to a growth inhibitory agent. It is

CC produced in bacteria or in CHO cells and induces death of a cell to which

CC it binds. The antibody is useful for preparing a composition for

CC diagnosing or treating tumours and cancer. The present sequence

CC represents a human TAT cDNA sequence of the invention.

XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	123	Length:	2053
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-10 (1-9) x ADL06473 (1-2053)

QY 1 PheLeuGlyTleValLeuAlaLeuile 9

DB 1156 TTCCTGGGTATTGTTTAGCCCTGATA 1182

RESULT 24

ADN03961

ID ADN03961 standard; cDNA; 2053 BP.

XX ADN03961;

XX 01-JUL-2004 (first entry)

XX Antipsoriatic cDNA sequence #180.

XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

XX WO2004028479-A2.

XX 08-APR-2004.

XX 25-SEP-2003; 2003WO-US030907.

XX 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

XX Wu TD;

XX WPI; 2004-305105/28.

XX P-PSDB; ADN03962.

XX New PRO nucleic acid or polypeptide, useful for preparing a

PT pharmaceutical composition for diagnosing or treating psoriasis in a

PT mammal.

XX Claim 1; SEQ ID NO 355; 3069pp; English.

CC The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polynucleotides of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 123 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-10 (1-9) x ADN03961 (1-2053)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 DB 1156 TTCTGGGTATTGTTTAGCCCTGATA 1182

RESULT 25
 ADR25444
 ID ADR25444 standard; DNA; 2053 BP.

XX AC ADR25444;
 XX AC
 XX 21-OCT-2004 (first entry)

XX DT Breast cancer prognosis marker #1305.

XX DE ds; breast cancer; prognosis; gene expression; diagnosis.

XX KW Homo sapiens.

XX OS WO2004065545-A2.

XX PN 05-AUG-2004.

XX PD 15-JAN-2004; 2004WO-US001100.

XX PF 15-JAN-2003; 2003US-00342887.

XX PR (ROSE-) ROSETTA INPHARMATICS LLC.
 XX PA (NECA-) NETHERLANDS CANCER INST.

XX PI Van't Veer LJ, He Y;

XX DR WPI; 2004-593473/57.

XX PT Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of
 PT five genes in a cell sample taken from patient, to control levels.

XX PS Disclosure; SEQ ID NO 1305; 226pp; English.

XX CC The invention relates to a method of classifying a breast cancer patient
 CC according to prognosis by determining the similarity between the level of
 CC expression of each of five genes for which markers are listed in the
 CC specification, in a cell sample taken from the breast cancer patient, to
 CC control levels of expression for each respective five genes to obtain a
 CC patient similarity value. The methods are useful for classifying a breast
 CC cancer patient according to prognosis. Kits and computer program products
 CC are useful for data analysis using the diagnostic, prognostic and
 CC statistical methods of the invention. This sequence corresponds to a
 CC marker used in the method of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 123 Length: 2053
 Score: 40.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-10 (1-9) x ADR25444 (1-2053)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 DB 1156 TTCTGGGTATTGTTTAGCCCTGATA 1182

RESULT 26
 ACN38510
 ID ACN38510 standard; cDNA; 2053 BP.

XX AC ACN38510;

XX AC 18-NOV-2004 (first entry)

XX DE Tumour-associated antigenic target (TAT) cDNA DNAL03471, SEQ ID NO:2070.

XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic; gene; ss.

XX OS Homo sapiens.

XX PN WO2004030615-A2.

XX PD 15-APR-2004.

XX PF 29-SEP-2003; 2003WO-US028547.

XX PR 02-OCT-2002; 2002US-0414971P.

XX PA (GETH) GENENTECH INC.

XX PI Wu TD, Zhang Z, Zhou Y;

XX DR WPI; 2004-347921/32.

XX DR P-PSDB; ABM80804.

XX PT New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.

XX PS Claim 1; SEQ ID NO 2070; 7273pp; English.

XX CC The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT nucleic acid of the invention

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 123 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-10 (1-9) x ACN38510 (1-2053)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 1156 TTCTGGGTATTTGTTAGCCCTGATA 1182

RESULT 27

ID ADV35098

ADV35098 standard; cDNA; 2053 BP.

AC ADV35098;

DT 10-FEB-2005 (first entry)

DE Human cDNA of an exemplary efficacy gene for BAD SeqID174.

XX human; ss; multi-parameter high throughput screening; MPHTS;

KW disease signature; neuropsychiatric; neurodegenerative; schizophrenia;

KW bipolar affective disorder; BAD; autism; Parkinson's;

KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.

XX Homo sapiens.

XX US2003096264-A1.

PN 22-MAY-2003.

PD 18-JUN-2002; 2002US-00175523.

PF 18-JUN-2001; 2001US-0299151P.

XX 07-SEP-2001; 2001US-0317828P.

PR 25-SEP-2001; 2001US-0325150P.

PR 14-NOV-2001; 2001US-0333047P.

PR 18-JAN-2002; 2002US-034936P.

PR 04-MAR-2002; 2002US-0361834P.

XX (PSYC-) PSYCHIATRIC GENOMICS INC.

XX Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;

PI Palfreyman M, Rajan P;

XX WPI; 2004-118903/12.

DR Identifying a compound that can treat disease or disorders, such as, a

PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises

PT determining the expression of one or more efficacy genes in a cell

PT contacted with the test compound.

XX Example 6; SEQ ID NO 174; 39pp; English.

XX This invention relates to a novel screening method identified as a multi-

CC parameter high throughput screening (MPHTS) assay. Specifically, it

CC refers to an assay that utilizes the disease signature of a plurality of

CC specific genes associated with a particular disease, and identifies

CC differential expression between those cells taken from individuals

CC affected by that disease and those that are not affected. The present

CC invention then describes the screening of candidate pharmaceutical

CC compounds to identify those that have a potential therapeutic benefit for

CC the treatment of neuropsychiatric and neurodegenerative disorders

CC including schizophrenia, bipolar affective disorder (BAD) and autism, as

CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds

CC of this invention exhibit various activities including neuroleptic,

CC nootropic, antimanic and antidepressant. Furthermore, the screening

CC method used in MPHTS will be automated, such that a large number of test

CC compounds may be rapidly screened with a minimal amount of labour and

CC effort. This polynucleotide is a human cDNA sequence of a gene that is

CC differentially expressed in the presence of a therapeutic compound and

CC represents an exemplary efficacy gene for bipolar affective disorder,

CC given in an exemplification of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 123 Length: 2053

Score: 40.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 13 Gaps: 0

US-10-774-176-10 (1-9) x ADV35098 (1-2053)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 1156 TTCTGGGTATTTGTTAGCCCTGATA 1182

RESULT 28

AAS87175

ID AAS87175 standard; cDNA; 2338 BP.

XX AAS87175;

AC AAS87175;

XX 13-FEB-2002 (first entry)

DT DNA encoding novel human diagnostic protein #22979.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

PR (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG22988.

DR New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 1; SEQ ID NO 22979; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 143 Length: 2338
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-10 (1-9) x AAS87175 (1-2338)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1413 TTCTGGGTATTGTTTAGCCCTGATA 1439

RESULT 29

AAK94253
ID AAK94253 standard; cDNA; 2359 BP.

XX AAK94253;

XX 06-NOV-2001 (first entry)

XX Human full-length cDNA, SEQ ID NO: 2864.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX P-PSDB; AAM93333.

XX Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.

XX This invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesizing the full length cDNA are useful
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a full length human cDNA of the

CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in CD-ROM format directly
CC from EPO

XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 144 Length: 2359
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-10 (1-9) x AAK94253 (1-2359)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1495 TTCTGGGTATTGTTTAGCCCTGATA 1521

RESULT 30

ADL30831

ID ADL30831 standard; cDNA; 2359 BP.

XX ADL30831;

XX 20-MAY-2004 (first entry)

XX Full length human cDNA clone SeqID 2864.

XX human; medicine; signal transduction; glycoprotein; transcription;
XX oligo-capping method; ss; gene.

XX Homo sapiens.

XX EP1396543-A2.

XX 10-MAR-2004.

XX 07-JUL-2000; 2003EP-00025638.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183865.

XX 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

XX P-PSDB; ADL30832.

XX Example 1; SEQ ID NO 2864; 1340pp; English.

XX This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.

XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 144 Length: 2359

Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-10 (1-9) x ADL30831 (1-2359)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 1495 TTCCTGGGTATTGTTTTCAGCCCTGATA 1521

RESULT 31

AAK94254

ID AAK94254 standard; cDNA; 2361 BP.

XX AAK94254;

XX 06-NOV-2001 (first entry)

XX Human full-length cDNA, SEQ ID NO: 2866.

DE Human, full length cDNA; cDNA synthesis; oligo-capping; ss.

KW Homo sapiens.

OS Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 12-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

DR P-PSDB; AAK93334.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

PT in genetic manipulation.

PT Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been isolated

CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have

CC been determined. Primers for synthesizing the full length cDNA are useful

CC for clarifying the function of the protein encoded by the cDNA. The full

CC length clones were obtained by construction of full length enriched cDNA

CC libraries that were synthesised by the oligo-capping method. The primers

CC enable the production of the full length cDNA easily without any special

CC methods. The present sequence is a full length human cDNA of the

CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in CD-ROM format directly

CC from EPO

XX SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 144 Length: 2361
 Score: 40.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-10 (1-9) x AAK94254 (1-2361)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
 Db 1497 TTCCTGGGTATTGTTTTCAGCCCTGATA 1523

RESULT 32

ADI26162

ID ADI26162 standard; cDNA; 2361 BP.

XX ADI26162;

XX 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #64.

XX ss; gene; human; signal transducer and activator of transcription 6;

KW STAT6; immunogen; STAT6 activation; allergy; inflammation; cancer;

KW autoimmune disease; diabetes; hyperlipidaemia; infection; rheumatoid arthritis; osteoarthritis;

KW Th1 hyperactive disease; rheumatoid arthritis; asthma; allergic rhinitis;

KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

OS Homo sapiens.

XX WO2003104277-A2.

XX 18-DEC-2003.

XX 05-JUN-2003; 2003WO-JP007123.

XX 05-JUN-2002; 2002JP-00164257.

PR 06-JUN-2002; 2002US-0385912P.

PR 26-DEC-2002; 2002JP-00377326.

PR 27-DEC-2002; 2002US-0436467P.

PR 15-MAY-2003; 2003JP-00137505.

PR 16-MAY-2003; 2003US-0470836P.

XX (ASAH) ASAH KASEI KK.

XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX WPI; 2004-122214/12.

DR P-PSDB; ADI26163.

XX New signal transducer and activator of transcription 6 activation

PT promoting purified protein, for diagnosing and treating disease

PT associated with activation/inhibition of transcription factor e.g.

PT diabetes and cancer.

XX Claim 4; SEQ ID NO 127; 1368pp; English.

XX The invention relates to a purified protein promoting signal transducer

CC and activator of transcription 6 activation (STAT6). The protein is

CC useful for the producing an antibody, which involves administering the

CC protein or its epitope-bearing fragments to a non-human animal as an

CC antigen. The nucleic acid is useful for diagnosing a disease or

CC susceptibility to a disease related to expression or activity of the

CC protein. A transformant expressing the protein is useful for screening

CC compounds which inhibit or promote STAT6 activation. A transformant

CC expressing the protein is useful for producing a pharmaceutical

CC composition. Compositions, antibodies and antisense molecules are useful

CC for the treating a disease associated with STAT6 activation such as

CC allergic diseases, inflammation, autoimmune diseases, diabetes,

CC hyperlipidaemia, infections disease and cancers. Compositions are useful

CC for treating disease associated with STAT6 activation and/or prevention

CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid

CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,

CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,

CC viral hepatitis and AIDS. The protein has effectively useful for screening

CC activity. The protein or nucleic acid is effectively useful for screening

CC compounds for treating and preventing disease associated with excessive

CC activation or inhibition of STAT6. The present sequence represents a

CC human cDNA encoding a protein which promotes STAT6 activation.

```
XX SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 144 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-10 (1-9) x ADL30833 (1-2361)
QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1497 TTCTGGGTATTGTTTAGCCCTGATA 1523

RESULT 34
ADI26160
ID ADI26160 standard; cDNA; 2557 BP.
XX
AC ADI26160;
XX
DT 22-APR-2004 (first entry)
XX
DE Human cDNA encoding protein that promotes STAT6 activation #63.
XX
KW ss; gene; human; signal transducer and activator of transcription 6;
KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
XX
OS Homo sapiens.
XX
FN WO2003104277-A2.
XX
PD 18-DEC-2003.
XX
PF 05-JUN-2003; 2003WO-JP007123.
XX
PR 05-JUN-2002; 2002JP-00164257.
PR 06-JUN-2002; 2002US-0385912P.
PR 26-DEC-2002; 2002JP-00377326.
PR 27-DEC-2002; 2002US-0436467P.
PR 15-MAY-2003; 2003JP-00137505.
PR 16-MAY-2003; 2003US-0470836P.
XX
PA (ASAH) ASahi KASEI KK.
XX
PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
XX
DR WPI; 2004-122214/12.
XX
P-PSDB; ADI26161.
XX
PT New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX
PS Claim 4; SEQ ID NO 125; 1368pp; English.
XX
CC The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infections disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
```

CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.

XX
SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 158 Length: 2557
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-10 (1-9) x ADI26160 (1-2557)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1645 TTCTAGGTATTGTTTTCAGCTCTGATA 1671
|||||

RESULT 35
ADI26158
ID ADI26158 standard; cDNA; 2557 BP.

XX
AC ADI26158;

XX
DT 22-APR-2004 (first entry)

XX
DE Human cDNA encoding protein that promotes STAT6 activation #62.

XX
KW ss; Gene; human; signal transducer and activator of transcription 6;
KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

XX
OS Homo sapiens.

XX
PN WO2003104277-A2.

XX
PD 18-DEC-2003.

XX
PF 05-JUN-2003; 2003WO-JP007123.

XX
PR 05-JUN-2002; 2002JP-00164257.

XX
PR 06-JUN-2002; 2002US-0385912P.

XX
PR 26-DEC-2002; 2002JP-00377326.

XX
PR 27-DEC-2002; 2002US-0436467P.

XX
PR 15-MAY-2003; 2003JP-00137505.

XX
PR 16-MAY-2003; 2003US-0470836P.

XX
PA (ASAH) ASAH KASEI KK.

XX
PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX
WPI; 2004-122214/12.

XX
P-PSDB; ADI26159.

XX
PT New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.

XX
PS Claim 4; SEQ ID NO 123; 1368pp; English.

XX
CC The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the

CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.

XX
SQ Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 158 Length: 2557
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-10 (1-9) x ADI26158 (1-2557)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1645 TTCTAGGTATTGTTTTCAGCTCTGATA 1671
|||||

RESULT 36

ADL03007/c

ID ADL03007 standard; DNA; 207 BP.

XX
AC ADL03007;

XX
DT 06-MAY-2004 (first entry)

XX
DE DNA encoding a M. catarrhalis protein #693.

XX
KW ds; gene; Moraxella catarrhalis; infection.

XX
OS Moraxella catarrhalis.

XX
PN US6673910-B1.

XX
PD 06-JAN-2004.

XX
PF 04-APR-2000; 2000US-00540236.

XX
PR 08-APR-1999; 99US-0128416P.

XX
PA (GENO-) GENOME THERAPEUTICS CORP.

XX
PI Breton GL;

XX
WPI; 2004-178127/17.

XX
P-PSDB; ADL04927.

XX
PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.

XX
PS Disclosure; SEQ ID NO 693; 429pp; English.

XX
CC The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents DNA encoding a M.
CC catarrhalis protein.

XX SQ Sequence 207 BP; 68 A; 30 C; 51 G; 58 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 64.8 Length: 207
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-10 (1-9) x ADL03007 (1-207)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 91 TTTCAGGCGCTGTGTAGCTCTCTT 65

RESULT 37
ABK79121
ID ABK79121 standard; DNA; 312 BP.
XX AC ABK79121;
XX DT 13-AUG-2002 (first entry)
XX DE Bacillus clausii genomic sequence tag (GST) #1964.
XX KW Differential gene expression; genomic sequenced tag; GST;
XX KW altered culture condition; environmental stress;
XX KW physiological provocation; ds.
XX OS Bacillus clausii.
XX PN WO200229113-A2.
XX PD 11-APR-2002.
XX PF 05-OCT-2001; 2001WO-US031437.
XX PR 06-OCT-2000; 2000US-00680598.
XX PR 27-MAR-2001; 2001US-0279526P.
XX PA (NOVO) NOVOZYMES BIOTECH INC.
XX PA (NOVO) NOVOZYMES AS.
XX PI Berka R, Clausen IG;
XX WPI; 2002-416684/44.
XX Monitoring differential expression of several genes in first Bacillus
cell relative to expression of same genes in one or more second Bacillus
cells, by using substrate containing Bacillus genomic sequenced tag
array.

Claim 11; SEQ ID NO 6412; 200pp; English.

The invention describes a method of monitoring differential expression of
genes in a first Bacillus cell relative to expression of the genes in
other Bacillus cells, comprising hybridising labelled nucleic acid probes
isolated from Bacillus cells to a substrate containing array of Bacillus
genomic sequenced tags (GST), examining the array, and determining
relative gene expression by an observed hybridisation reporter signal of
a spot in the array. The method is useful for measuring the expression of
genes in a first Bacillus cell relative to expression of the same genes
in one or more second Bacillus cells. The method is useful for monitoring
global expression of several genes from a Bacillus cell, discovering new
genes, identifying possible functions of unknown open reading frames and
monitoring gene copy number variation and stability. Monitoring changes
in expression of genes may be used to provide a representation of the way
in which Bacillus cells adapt to changes in culture conditions,
environmental stress or other physiological provocation. Extensive follow
-up characterisation is unnecessary, when one spot on an array equals one
gene or one open reading frame, since sequence information is available.

CC This sequence represents a genomic sequence tag (GST) used in the method
of the invention. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 312 BP; 76 A; 62 C; 65 G; 109 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 104 Length: 312
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x ABK79121 (1-312)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 179 TTCATTGGATCGTGTGCTTAATT 205

RESULT 38
AAS37087/c
ID AAS37087 standard; cDNA; 402 BP.
XX AC AAS37087;
XX DT 17-DEC-2001 (first entry)
XX DE Novel human diagnostic and therapeutic gene #145.
XX KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.
XX OS Homo sapiens.
XX PN WO200166753-A2.
XX PD 13-SEP-2001.
XX PF 09-MAR-2001; 2001WO-US007787.
XX PR 09-MAR-2000; 2000US-0188609P.
XX PA (CHIR) CHIRON CORP.
XX PA (HYSE-) HYSEQ INC.
XX PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;
XX WPI; 2001-530177/58.
XX New polynucleotides and polypeptides, useful for diagnosis and treatment
of breast, lung and colon cancer.
XX Claim 1; Page 632; 1193pp; English.
XX The invention relates to new polynucleotides and polypeptides, useful for
diagnosis and treatment of breast, lung and colon cancer. The sequences
can be used in detecting differentially expressed genes correlated with a
cancerous state of a mammalian cell, comprising detecting at least one
differentially expressed gene product in a test sample derived from a
cell suspected of being cancerous. They can also be used to inhibit
tumour growth by modulating expression of a gene product. AAS36943-
CC AAS39338 represent novel human diagnostic and therapeutic coding
sequences of the invention
XX SQ Sequence 402 BP; 91 A; 92 C; 110 G; 109 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 140 Length: 402
Score: 36.00 Matches: 7

Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	90.0%	Indels:	0
DB:	10	Gaps:	0

US-10-774-176-10 (1-9) x ADK53767 (1-575)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 413 TTCTTGGAGTTGTAGTCTCTCATA 387

RESULT 40
AAL19859
ID AAL19859 standard; cDNA; 689 BP.
XX
XX AAL19859;
XX
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 12316.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
PP 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
XX 14-MAR-2000; 2000US-0189167P.
XX 24-MAR-2000; 2000US-0192099P.
XX 29-MAR-2000; 2000US-0193480P.
XX 15-MAY-2000; 2000US-0205230P.
XX 09-JUN-2000; 2000US-0211315P.
XX 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX Claim 1; Page 2176; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterizing treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

SQ Sequence 689 BP; 219 A; 99 C; 116 G; 255 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	261	Length:	689
Score:	36.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	90.0%	Indels:	0
DB:	4	Gaps:	0

US-10-774-176-10 (1-9) x AAL19859 (1-689)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 413 TTCTTGGAGTTGTAGTCTCTCATA 387

Db 172 TTTTGGGTTTGGTGGCAATTATT 198

RESULT 41

ABK79169

ID ABK79169 standard; DNA; 829 BP.

AC ABK79169;

XX

XX

DT 13-AUG-2002 (first entry)

XX

DE Bacillus clausii genomic sequence tag (GST) #2012.

XX

XX Differential gene expression; genomic sequenced tag; GST;

KW altered culture condition; environmental stress;

KW physiological provocation; ds.

XX

OS Bacillus clausii.

XX

PN WO200229113-A2.

XX

XX 11-APR-2002.

XX

XX

PF 05-OCT-2001; 2001WO-US031437.

XX

PR 06-OCT-2000; 2000US-00680598.

PR

PR 27-MAR-2001; 2001US-0279526P.

XX

XX (NOVO) NOVOZYMES BIOTECH INC.

PA (NOVO) NOVOZYMES AS.

PA

XX Berka R, Clausen IG;

XX

XX WPI; 2002-416684/44.

DR

XX

XX Monitoring differential expression of several genes in first Bacillus

PT cell relative to expression of same genes in one or more second Bacillus

PT cells, by using substrate containing Bacillus genomic sequenced tag

PT array.

XX

PS Claim 11; SEQ ID NO 6460; 200pp; English.

XX

XX The invention describes a method of monitoring differential expression of

CC genes in a first Bacillus cell relative to expression of the genes in

CC other Bacillus cells, comprising hybridising labelled nucleic acid probes

CC isolated from Bacillus cells to a substrate containing array of Bacillus

CC genomic sequenced tags (GST), examining the array, and determining

CC relative gene expression by an observed hybridisation reporter signal of

CC a spot in the array. The method is useful for measuring the expression of

CC genes in a first Bacillus cell relative to expression of the same genes

CC in one or more second Bacillus cells. The method is useful for monitoring

CC global expression of several genes from a Bacillus cell, discovering new

CC genes, identifying possible functions of unknown open reading frames and

CC monitoring gene copy number variation and stability. Monitoring changes

CC in expression of genes may be used to provide a representation of the way

CC in which Bacillus cells adapt to changes in culture conditions,

CC environmental stress or other physiological provocation. Extensive follow

CC -up characterisation is unnecessary, when one spot on an array equals one

CC gene or one open reading frame, since sequence information is available.

CC This sequence represents a genomic sequence tag (GST) used in the method

CC of the invention. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 829 BP; 204 A; 160 C; 171 G; 294 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 324 Length: 829

Score: 36.00 Matches: 7

Percent Similarity: 100.0% Conservatives: 2

Best Local Similarity: 77.8% Mismatches: 0

Query Match: 90.0% Indels: 0

DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x ABK79169 (1-829)

Qy 1 PheLeuglyfIleValLeuAlaIeulle 9

Db 658 TTCAATGGATCGTGAATTCCTTAATT 684

RESULT 42

ACA39024/C

ID ACA39024 standard; DNA; 1833 BP.

XX

XX ACA39024;

XX

DT 19-JUN-2003 (first entry)

XX

DE Prokaryotic essential gene #20681.

XX

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

XX

OS Moraxella catarrhalis.

XX

XX WO200277183-A2.

PN

XX

PD 03-OCT-2002.

XX

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

XX

PR 21-MAR-2001; 2001US-00815242.

PR

PR 06-SEP-2001; 2001US-00948993.

PR

PR 25-OCT-2001; 2001US-0342923P.

PR

PR 08-FEB-2002; 2002US-00072851.

PR

PR 06-FEB-2002; 2002US-0362699P.

PR

XX

XX (ELIT-) ELITRA PHARM INC.

PA

XX

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI; 2003-029926/02.

DR

DR P-PSDB; ABU35154.

XX

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 14; SEQ ID NO 26894; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway (8)

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1833 BP; 538 A; 368 C; 424 G; 503 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 812 Length: 1833
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x ACA39024 (1-1833)
Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
Db 1720 TTCTAGGCTGTTTCTAGCTCTCTT 1694

RESULT 43
ACA211440
ID ACA211440 standard; DNA; 2271 BP.
XX ACA211440;
XX
XX
XX 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #3097.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; Gene.
XX
XX Bacillus anthracis.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US0009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX F-PSDB; ABU17570.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 9310; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2271 BP; 826 A; 337 C; 537 G; 571 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.04e+03 Length: 2271
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x ACA21440 (1-2271)
Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
Db 838 TTGTAGGATTTGTTCTTCTTGT 864

RESULT 44
AAF28548/C
ID AAF28548 standard; DNA; 96109 BP.
XX
XX AAF28548;
XX
XX 04-APR-2001 (first entry)
XX
XX Genomic fragment #35.
XX
XX Genomic library; bacteria; human upper airway; otitis media; sinusitis;
XX bronchopulmonary; endocarditis; meningitis; ss.
XX
XX Moraxella catarrhalis.
XX
XX WO200078968-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US016649.
XX
XX 18-JUN-1999; 99US-0140121P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Lagace RE, Patterson C, Berg KL;
XX WPI; 2001-041427/05.
XX
XX Genomic library for identifying diagnostic and therapeutic compositions,
XX and for identifying virulence factors, regulatory elements and drug
XX targets, comprises Moraxella catarrhalis nucleic acids.
XX
XX Claim 1; Page 345-368; 545pp; English.
XX

CC The present invention relates to a Moraxella catarrhalis genomic library
 CC comprising of a combination of 41 nucleic acid molecules (see AAF28514-
 CC AAF28554). The library has a number of uses described in the
 CC specification e.g. is useful for identifying diagnostic and therapeutic
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
 CC aerobic, gram-negative diplococcus, normally found among the bacterial
 CC flora of human upper airways. M. catarrhalis is known to cause acute,
 CC localised infections such as otitis media, sinusitis and bronchopulmonary
 CC infection and life-threatening, systemic diseases including endocarditis
 CC and meningitis

XX SQ Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 8.01e+04 Length: 96109
 Score: 36.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 90.0% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-10 (1-9) x AAF28548 (1-96109)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 |||||:::|||||:::
 DB 4034 TTCTAGGGCTTGTTCCTCTCTT 4008

RESULT 45

ABD33266

ID ABD33266 standard; DNA; 168407 BP.

XX AC ABD33266;

XX DT 18-NOV-2004 (first entry)

XX DE Murine cancer-associated (CA) gene MD07-046.

XX KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW de; cancer; cytostatic.
 XX OS Mus musculus.
 XX PN W02004058146-A2.
 XX PD 15-JUL-2004.
 XX PF 15-DEC-2003; 2003WO-US040081.
 XX PR 17-DEC-2002; 2002US-00322281.
 XX (SAGR-) SAGRES DISCOVERY INC.
 XX PA Morris DW, Malandro MS;
 XX WPI; 2004-499109/47.

XX Novel human cancer associated protein encoded within open reading frame
 PT of cancer associated gene, useful as targets for diagnosing cancer.
 XX Disclosure; SEQ ID NO 305; 182pp; English.
 XX The invention relates to cancer-associated proteins (CAP) and the cancer-
 CC associated (CA) nucleic acids encoding them. The invention also relates
 CC to a method for treating cancers involving administering to a patient an
 CC inhibitor of CAP, and a method of screening for anticancer activity in a
 CC potential drug involving providing a cell that expresses a CA gene,
 CC contacting a tissue sample derived from a cancer cell with an anticancer
 CC drug candidate and monitoring the effect of the anticancer drug candidate
 CC on expression of the CA gene. The CAP proteins are useful for detecting
 CC cancer associated with expression of a CAP protein in a test cell sample
 CC and for screening for a bioactive agent capable of modulating the
 CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 CC cancer, involving determining the expression of a CA nucleic acid in a

CC tissue. This sequence represents a murine CA gene of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 168407 BP; 46388 A; 34421 C; 36717 G; 50689 T; 0 U; 192 Other;

Alignment Scores:
 Pred. No.: 1.52e+05 Length: 168407
 Score: 36.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 90.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-10 (1-9) x ABD33266 (1-168407)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
 |||||:::|||||:::
 DB 18784 TTTTAGGTATAGTAGTGGCTCTTCTT 18810

RESULT 46

ABQ81842/C

ID ABQ81842 standard; DNA; 349980 BP.

XX AC ABQ81842;

XX DT 19-NOV-2002 (first entry)

XX DE Bifidobacterium longum NCC2705 genomic sequence SEQ ID NO:1.

XX KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
 KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
 KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
 KW rotavirus; food composition; pharmaceutical composition; gene; ds.
 XX OS Bifidobacterium longum.

XX EP1227152-A1.

XX PD 31-JUL-2002.

XX PF 30-JAN-2001; 2001EP-00102050.

XX PR 30-JAN-2001; 2001EP-00102050.

XX PA (NEST) SOC PROD NESTLE SA.

XX WPI; 2002-668397/72.

XX Novel polynucleotide comprising Bifidobacterium genome sequence useful as
 PT a probe or primer for detecting and/or identifying Bifidobacterium longum
 PT in a biological sample.
 XX Claim 1; SEQ ID NO 1; 80pp; English.
 XX The present invention describes a polynucleotide (I) comprising a
 CC sequence of a Bifidobacterium genome selected from the nucleotide
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
 CC least 90% identity or which hybridises with the sequences given in
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
 CC fusion protein, comprising a sequence selected from 1097 sequences given
 CC in ABP65258 to ABP65354 ligated in frame to a polynucleotide encoding a
 CC heterologous polypeptide. (I) has antidiarrheic and antibacterial
 CC activities, and can be used as an inhibitor of Salmonella. (I) (which is
 CC a probe) is useful for the detection and/or identification of
 CC Bifidobacterium longum in a biological sample. A carrier containing the
 CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be
 CC used for preventing and/or treating diarrhoea brought about by pathogenic
 CC bacteria and/or rotavirus. The carrier is a food composition selected
 CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
 CC products, ice-creams, fermented cereal based products, milk based
 CC powders, infant formula, pet food or a pharmaceutical composition

CC selected from tablets, liquid bacterial suspensions, dried oral
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC (I) is useful in DNA arrays or chips to carry out analysis of the
CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
CC Bifidobacterium related nucleotide sequences given in the Sequence
CC Listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office

XX
XX
SQ Sequence 349980 BP; 72540 A; 102738 C; 103221 G; 71481 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.52e+05 Length: 349980
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x ABQ81844 (1-349980)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 147704 TTCTTGGCCTTCTTCTGCCCTTCTC 147678

RESULT 47
ACF73669
ID ACF73669 standard; DNA; 666 BP.
XX
AC ACF73669;
XX
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus DNA #1349.
DE
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target; gene; ds.
XX
XX Staphylococcus aureus.

XX WO200294868-A2.
XX
XX 28-NOV-2002.
XX
XX 27-MAR-2002; 2002WO-IB002637.
XX
XX 27-MAR-2001; 2001GB-00007661.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Masignani V, Mora M, Scarselli M;
XX
XX WPI; 2003-120786/11.
DR P-PSDB; ABM72109.
XX
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX
XX Claim 6; SEQ ID NO 2697; 49pp; English.

XX The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus genes of the invention

SQ Sequence 666 BP; 257 A; 90 C; 110 G; 209 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 416 Length: 666
Score: 35.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 87.5% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x ACF73669 (1-666)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 538 TTTTAGGATTAGTGTGCGCTTATA 564

RESULT 48
ABS63032/C
ID ABS63032 standard; DNA; 853 BP.
XX
XX ABS63032;
XX
XX 05-NOV-2002 (first entry)

XX Selected Interacting Domain (SID) polynucleotide #229.
DT
DE
XX Yeast; selected interacting domain; SID; gene; ds; antifungal; cancer;
KW cytostatic; neuroprotective; Candida infection; gene therapy;
KW neurodegenerative disease.
XX
XX Saccharomyces cerevisiae.

XX WO200259255-A2.
XX
XX 01-AUG-2002.
XX
XX 25-JAN-2002; 2002WO-EP001350.
XX
XX 26-JAN-2001; 2001US-0264577P.
XX
XX (HYBR-) HYBRIGENICS.
XX
XX Legrain P;
XX
XX WPI; 2002-619165/66.
DR P-PSDB; ABG77418.
XX
XX New complex between two interacting bait and prey Saccharomyces
PT cerevisiae polypeptides, useful for preventing or treating Candida
PT infection, cancer or neurodegenerative diseases in a mammal.
XX
XX Claim 7; Page 158; 196pp; English.

XX The invention relates to a complex between two interacting Saccharomyces
CC cerevisiae polypeptides, comprising two Selected Interacting Domain (SID)
CC polypeptides as bait and prey proteins. A pharmaceutical composition
CC comprising the complex is useful for preventing or treating Candida
CC infection, cancer and neurodegenerative diseases in a human or animal,
CC preferably in a mammal. This sequence represents DNA encoding a SID
CC polypeptide of the invention

SQ Sequence 853 BP; 309 A; 140 C; 152 G; 252 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 555 Length: 853
Score: 35.00 Matches: 6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match: 87.5% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x ABS63032 (1-853)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 425 TTCTTCGAGTAGTCTTAAGTTAGTG 399

RESULT 49
ABT11637/c
ID ABT11637 standard; DNA; 853 BP.
XX
XX
AC ABT11637;
XX
XX
DT 10-DEC-2002 (first entry)
XX
DE Yeast selected interacting domain coding sequence SEQ ID NO: 773.
XX
KW Yeast; protein-protein interaction; Selected Interacting Domain;
KW SID (RTM); secretion yield; cancer; neurodegenerative disease; fungicide;
KW cytosolic; neuroprotective; gene; ds.
XX
XX Saccharomyces cerevisiae.
XX
XX WO200266504-A2.
XX
XX 29-AUG-2002.
XX
XX 14-FEB-2002; 2002WO-EP002299.
XX
XX 16-FEB-2001; 2001US-0269266P.
XX
XX (HYBR-) HYBRIGENICS.
XX
XX Legrain P;
XX
XX WPI; 2002-674913/72.
XX
XX P-PSDB; ABJ11320.
XX
XX New protein-protein complexes of Saccharomyces cerevisiae, useful in drug
PT screening or development, for developing yeast strains with better
PT secretion yield of protein, or in gene therapy (e.g. to treat Candida
PT infection or cancer).
XX
XX Claim 7; Page 310; 357pp; English.
XX
XX The present invention relates to complexes between Saccharomyces
CC cerevisiae Selected Interacting Domain (SID (RTM)) proteins and coding
CC sequences. The protein complexes of S. cerevisiae are useful in drug
CC development, in screening drugs or agents that modulate the interaction
CC of proteins, for developing yeast strains with better secretion yield of
CC protein, and in gene therapy. The protein complexes, polypeptides and
CC polynucleotides are useful for preventing or treating Candida infection,
CC cancer or neurodegenerative diseases in humans or animals. The present
CC sequence is a coding sequence of the invention
XX
SQ Sequence 853 BP; 309 A; 140 C; 152 G; 252 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 555 Length: 853
Score: 35.00 Matches: 6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match: 87.5% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x ABT11637 (1-853)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 425 TTCTTCGAGTAGTCTTAAGTTAGTG 399

RESULT 50
AAC37238
ID AAC37238 standard; DNA; 1326 BP.
XX
XX AAC37238;
AC

XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 16667.
XX
XX Arabidopsis thaliana
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search using frame_plus_p2n_model

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignment)
171.290 Million cell updates/sec

Title: US-10-774-176-10

Perfect score: 40

Sequence: 1 FLGIVLALI 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abes/ABSSWEB.spool/US1074176/runat_24042006_165114_19197/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes04
-USER=US1074176 @CGN 1.1 6765 @runat_24042006_165114_19197 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sy.*
12: gb.un.*
13: gb.vi.*
14: gb.htg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	290	6	CQ687716
2	40	100.0	475	6	CQ920916
3	40	100.0	927	6	AX829164

4	40	100.0	1260	6	AX467373
5	40	100.0	1260	6	AX821533
6	40	100.0	1260	6	AX821548
7	40	100.0	1263	6	BD249731
8	40	100.0	1263	6	AX025011
9	40	100.0	1263	6	AX149553
10	40	100.0	1263	6	AX316086
11	40	100.0	1263	6	AX467371
12	40	100.0	1281	6	BD249732
13	40	100.0	1281	6	AX025012
14	40	100.0	1281	6	AX316087
15	40	100.0	2053	6	CQ731678
16	40	100.0	2053	8	HS5T40A
17	40	100.0	2183	5	CR855786
18	40	100.0	2333	5	AF063939
19	40	100.0	2359	6	BD127282
20	40	100.0	2359	6	CQ782724
21	40	100.0	2359	8	AK074786
22	40	100.0	2361	6	BD127283
23	40	100.0	2361	6	CQ782726
24	40	100.0	2361	6	AX961916
25	40	100.0	2361	8	AK074790
26	40	100.0	2361	9	BC087011
27	40	100.0	2379	8	BC037161
28	40	100.0	2423	9	BC058198
29	40	100.0	2557	6	AX961912
30	40	100.0	2557	6	AX961914
31	40	100.0	2714	8	AB168308
32	40	100.0	5551	8	HSA012159
33	40	100.0	7942	9	MMU012160
34	40	100.0	121909	8	HSJ492P14
35	40	100.0	144888	5	BX248521
36	40	100.0	167046	9	AC158516
37	40	100.0	210237	14	AC128294
38	40	100.0	23076	14	AC106962
39	39	97.5	66940	14	AC165904
40	39	97.5	174454	14	AC150616
41	38	95.0	421	10	HSPA32B9
42	38	95.0	440	10	AU026869
43	38	95.0	58433	8	AL807766
44	38	95.0	151552	5	AL954182
45	38	95.0	164218	5	EX005450
46	38	95.0	164748	5	BX233031
47	38	95.0	166117	14	CR855265
48	38	95.0	213305	14	AC114043
49	38	95.0	232746	14	AC125778
50	38	95.0	271932	14	AC097673
51	37	92.5	388	10	HUMUT760A
52	37	92.5	1027	10	CNS06F52
53	37	92.5	13274	1	AE006847
54	37	92.5	98261	8	HS1054C24
55	37	92.5	125123	8	AC104455
56	37	92.5	147503	5	AC147906
57	37	92.5	155477	9	AC122752
58	37	92.5	161505	9	AC154234
59	37	92.5	162045	5	CR318660
60	37	92.5	164789	8	AC147185
61	37	92.5	171348	8	AF207955
62	37	92.5	177172	9	AC133652
63	37	92.5	181443	14	AC123326
64	37	92.5	184554	8	BS000168
65	37	92.5	184989	9	AC131652
66	37	92.5	189554	5	BX004812
67	37	92.5	190632	9	AC110622
68	37	92.5	191275	14	AC133067
69	37	92.5	192013	14	BX511255
70	37	92.5	195797	14	AC128145
71	37	92.5	196083	14	AC027092
72	37	92.5	197605	14	AC026905
73	37	92.5	205206	14	AC162296
74	37	92.5	210068	14	CT009510
75	37	92.5	219036	14	AC105588
76	37	92.5	221614	14	AC130976

C	77	37	92.5	231080	14	AC106214	AC106214 Rattus no	150	36	90.0	133785	14	CR936542	CR936542 Danio rer
	78	37	92.5	243444	14	AC133748	Rattus no	C 151	36	90.0	195041	9	AC125195	AC125195 Mus muscu
	79	37	92.5	250430	14	AC133745	Rattus no	C 152	36	90.0	195234	14	CR735105	CR735105 Danio rer
	80	37	92.5	272171	14	AC130848	Rattus no	C 153	36	90.0	196427	9	AC118017	AC118017 Mus muscu
	81	37	92.5	275741	14	AC123452	Rattus no	C 154	36	90.0	200353	9	AC166575	AC166575 Mus muscu
	82	37	92.5	340000	8	HS21C048	Sequence	C 155	36	90.0	204012	9	AL732613	AL732613 Mouse DNA
	83	36	90.0	207	6	AR450029	Sequence	C 156	36	90.0	206309	9	AL773582	AL773582 Mouse DNA
	84	36	90.0	312	6	AX437997	Sequence	C 157	36	90.0	211034	9	AL731709	AL731709 Mouse DNA
	85	36	90.0	402	6	AX245215	Sequence	C 158	36	90.0	211313	14	AC130524	AC130524 Rattus no
	86	36	90.0	588	10	BV310021	S236P6102	C 159	36	90.0	217060	5	BX324190	BX324190 Zebrafish
	87	36	90.0	689	6	CQ427292	Sequence	C 160	36	90.0	219318	14	AC137433	AC137433 Rattus no
	88	36	90.0	829	6	AX438045	Sequence	C 161	36	90.0	222121	9	AC012294	AC012294 Mus muscu
	89	36	90.0	1138	9	AF057139	Microtus	C 162	36	90.0	222895	14	AC084065	AC084065 Mus muscu
	90	36	90.0	1591	9	BC026642	Mus muscu	C 163	36	90.0	224522	14	AC106627	AC106627 Rattus no
	91	36	90.0	10462	1	AE004932	Pseudomon	C 164	36	90.0	230109	14	AC114532	AC114532 Rattus no
	92	36	90.0	16989	15	SPBC1988	AL023594 S.pombe c	C 165	36	90.0	230448	14	AC128463	AC128463 Rattus no
	93	36	90.0	20284	8	AL359876	Human DNA	C 166	36	90.0	235090	14	AC103148	AC103148 Rattus no
	94	36	90.0	58068	8	AL309350	AL359876 Homo sapi	C 167	36	90.0	239426	14	AC134476	AC134476 Rattus no
	95	36	90.0	72064	14	AC104421	AL359876 Homo sapi	C 168	36	90.0	240425	9	CNS07YOT	AL713839 Mus muscu
	96	36	90.0	80589	5	BX470257	CR933839 Zebrafish	C 169	36	90.0	240461	14	AC094416	AC094416 Rattus no
	97	36	90.0	88824	5	CR933839	CR933839 Zebrafish	C 170	36	90.0	241179	9	AC111141	AC111141 Mus muscu
	98	36	90.0	92651	14	AC162476	AL352476 Bos tauru	C 171	36	90.0	245916	14	AC103546	AC103546 Rattus no
	99	36	90.0	96109	6	AR408756	AR408756 Sequence	C 172	36	90.0	246710	14	AC094240	AC094240 Rattus no
	100	36	90.0	96109	6	AX067460	AX067460 Sequence	C 173	36	90.0	249825	9	AC154353	AC154353 Mus muscu
	101	36	90.0	110000	1	AE005674	Continuation (32 o	C 174	36	90.0	250257	14	AC114355	AC114355 Rattus no
	102	36	90.0	110000	1	AE014295	Continuation (2 of	C 175	36	90.0	261302	14	AC157227	AC157227 Bos tauru
	103	36	90.0	110000	1	AE017225	Continuation (10 o	C 176	36	90.0	263253	14	AC097819	AC097819 Rattus no
	104	36	90.0	110000	1	AE017225	Continuation (11 o	C 177	36	90.0	269832	14	AC112751	AC112751 Rattus no
	105	36	90.0	110000	1	AE017261	Continuation (7 of	C 178	36	90.0	269956	14	AC094934	AC094934 Rattus no
	106	36	90.0	110000	1	AE017308	AE017308 Mycoplasma	C 179	36	90.0	272329	14	AC094789	AC094789 Rattus no
	107	36	90.0	110000	1	AE017334	Continuation (10 o	C 180	36	90.0	276459	14	AC162970	AC162970 Bos tauru
	108	36	90.0	110000	1	AE017334	Continuation (11 o	C 181	36	90.0	287833	1	AE017267	AE017267 Bacillus
	109	36	90.0	110000	1	AE017355	Continuation (11 o	C 182	36	90.0	290029	1	AE017027	AE017027 Bacillus
	110	36	90.0	110000	1	AP066627	Continuation (13 o	C 183	36	90.0	292906	1	AE016988	AE016988 Shigella
	111	36	90.0	110000	1	BA000012	Continuation (56 o	C 184	36	90.0	304708	1	AE017001	AE017001 Bacillus
	112	36	90.0	110000	1	BA000012	Continuation (57 o	C 185	36	90.0	307511	14	AC119644	AC119644 Rattus no
	113	36	90.0	110000	1	BX571965	Continuation (5 of	C 186	36	90.0	310114	14	AC163053	AC163053 Bos tauru
	114	36	90.0	110000	1	CP000010	CP000010 Burkholde	C 187	36	90.0	339168	14	AC078987	AC078987 Homo sapi
	115	36	90.0	110000	1	CP000083	Continuation (46 o	C 188	36	90.0	349980	6	AX491683	AX491683 Sequence
	116	36	90.0	110000	14	AC096086	Continuation (3 of	C 189	36	90.0	349980	6	AX553947	AX553947 Sequence
	117	36	90.0	110000	15	AP008207	Continuation (412	C 190	35	87.5	375	13	AE079513	AE079513 Chrysanth
	118	36	90.0	129489	9	AL713860	AL713860 Mouse DNA	C 191	35	87.5	441	10	BV188268	BV188268 sqm15777
	119	36	90.0	132008	14	AC128151	AC128151 Rattus no	C 192	35	87.5	549	10	G87081	G87081 S209P6207RC
	120	36	90.0	132912	5	CR457442	CR457442 Zebrafish	C 193	35	87.5	573	10	BV405646	BV405646 S229P6257
	121	36	90.0	137468	15	AP003269	AP003269 Oryza sat	C 194	35	87.5	660	10	PM9E4B	PM9E4B Penicilli
	122	36	90.0	143148	9	AL607034	AL607034 Mouse DNA	C 195	35	87.5	666	6	AX619734	AX619734 Sequence
	123	36	90.0	151863	5	BX255906	BX255906 Zebrafish	C 196	35	87.5	687	10	PM1E11B	PM1E11B Penicilli
	124	36	90.0	154177	9	AC148988	AC148988 Mus muscu	C 197	35	87.5	740	9	AF337757	AF337757 Neotoma f
	125	36	90.0	156362	14	AC158460	AC158460 Rhinolph	C 198	35	87.5	740	9	AF337760	AF337760 Neotoma f
	126	36	90.0	159797	8	AC104697	AC104697 Homo sapi	C 199	35	87.5	750	2	AK174110	AK174110 Ciona int
	127	36	90.0	161495	14	CR855387	CR855387 Danio rer	C 200	35	87.5	933	15	AY072500	AY072500 Arabidops
	128	36	90.0	162309	5	AL954709	AL954709 Zebrafish	C 201	35	87.5	1032	15	AY042848	AY042848 Arabidops
	129	36	90.0	164317	8	AL390029	AL390029 Human DNA	C 202	35	87.5	1075	15	AY096531	AY096531 Arabidops
	130	36	90.0	165243	5	AL929124	AL929124 Zebrafish	C 203	35	87.5	1143	9	AF186821	AF186821 Neotoma m
	131	36	90.0	167475	9	AC090881	AC090881 Mus Muscu	C 204	35	87.5	1143	9	AF294345	AF294345 Neotoma m
	132	36	90.0	168030	14	AC091462	AC091462 Mus muscu	C 205	35	87.5	1143	9	AF298841	AF298841 Neotoma m
	133	36	90.0	168421	9	AC133328	AC133328 Mus muscu	C 206	35	87.5	1143	9	AF298846	AF298846 Neotoma m
	134	36	90.0	169865	8	AC079967	AC079967 Homo sapi	C 207	35	87.5	1143	9	AF298847	AF298847 Neotoma m
	135	36	90.0	170632	8	AL136321	AL136321 Human DNA	C 208	35	87.5	1143	9	AF298848	AF298848 Neotoma m
	136	36	90.0	174009	9	AC119943	AC119943 Mus muscu	C 209	35	87.5	1143	9	AF298849	AF298849 Neotoma m
	137	36	90.0	175860	14	AC123382	AC123382 Rattus no	C 210	35	87.5	1143	9	AF298849	AF298849 Neotoma m
	138	36	90.0	179099	9	AC090046	AC090046 Mus muscu	C 211	35	87.5	1143	9	AF376478	AF376478 Neotoma m
	139	36	90.0	179891	9	AC124603	AC124603 Mus muscu	C 212	35	87.5	1240	5	ZEPTRANA	L33470 Danio rerio
	140	36	90.0	180289	14	AC137249	AC137249 Rattus no	C 213	35	87.5	1241	5	ZEPTRAND	L33473 Danio rerio
	141	36	90.0	180944	14	AC073824	AC073824 Mus muscu	C 214	35	87.5	1326	15	AY086571	AY086571 Arabidops
	142	36	90.0	186279	9	AC121977	AC121977 Mus muscu	C 215	35	87.5	1358	15	AY065231	AY065231 Arabidops
	143	36	90.0	187100	9	AC164153	AC164153 Mus muscu	C 216	35	87.5	2228	15	SCU76688	U76688 Schizophyll
	144	36	90.0	188335	9	AC118733	AC118733 Mus muscu	C 217	35	87.5	3584	6	AX833621	AX833621 Sequence
	145	36	90.0	190308	9	AC164626	AC164626 Mus muscu	C 218	35	87.5	3584	8	AK095511	AK095511 Homo sapi
	146	36	90.0	190318	14	AC153204	AC153204 Bos tauru	C 219	35	87.5	3855	5	AJ720777	AJ720777 Gallus ga
	147	36	90.0	190623	9	AC113061	AC113061 Mus muscu	C 220	35	87.5	4217	15	YSCXR52X	L22856 Saccharomyc
	148	36	90.0	192182	5	BX649602	BX649602 Zebrafish	C 221	35	87.5	4242	9	GPIHKAAS	D21854 Cavia porce
	149	36	90.0	193352	14	BX294664	BX294664 Mus muscu	C 222	35	87.5	5738	6	AR353966	AR353966 Sequence

223	35	87.5	5738	6	AR535522	AR535522 Sequence	c 296	35	87.5	120194	8	AL355474	AL355474 Human DNA
224	35	87.5	6925	8	BC075797	BC075797 Homo sapi	c 297	35	87.5	124041	8	CNS01DSJ	AL121808 Human chr
225	35	87.5	11334	6	AX346135	AX346135 Sequence	c 298	35	87.5	124189	15	AC146630	AC146630 Medicago
226	35	87.5	12278	13	HCU45478	U45478 Hog cholera	c 299	35	87.5	127026	15	CNS08C81	AL731752 Oryza sat
227	35	87.5	12289	13	AY259122	AY259122 Classical	c 300	35	87.5	130626	8	HS1114G22	AL008626 Human DNA
228	35	87.5	12297	13	AF091661	AF091661 Hog chole	c 301	35	87.5	131707	5	BX247948	BX247948 Zebrafish
229	35	87.5	12297	13	HCVC0MSEQ	X96550 Hog cholera	c 302	35	87.5	133529	8	AC073416	AC073416 Homo sapi
230	35	87.5	12298	13	HCU45477	U45477 Hog cholera	c 303	35	87.5	135183	8	AL603703	AL603703 Human DNA
231	35	87.5	12298	13	HCVC0MGEN	X87939 Classical s	c 304	35	87.5	136155	15	CNS08C80	AL731751 Oryza sat
232	35	87.5	12298	13	HCVPOLYP1	D49532 Hog cholera	c 305	35	87.5	139338	9	AC131781	AC131781 Mus muscu
233	35	87.5	12298	13	HCVPOLYP2	D49533 Hog cholera	c 306	35	87.5	139446	14	AC151437	AC151437 Carollia
234	35	87.5	12301	6	CQ867021	CQ867021 Sequence	c 307	35	87.5	140167	8	AC004010	AC004010 Homo sapi
235	35	87.5	12301	11	AF326963	AF326963 Classical	c 308	35	87.5	140195	8	AC087163	AC087163 Homo sapi
236	35	87.5	12310	13	AF091507	AF091507 Hog chole	c 309	35	87.5	141512	14	CR457456	CR457456 Danio rer
237	35	87.5	12310	13	AF091507	AF091507 Hog chole	c 310	35	87.5	141918	14	AC150160	AC150160 Gallus ga
238	35	87.5	12310	13	AF531433	AF531433 Classical	c 311	35	87.5	143993	14	AC134835	AC134835 Mus muscu
239	35	87.5	12311	6	AY382481	AY382481 Classical	c 312	35	87.5	148202	9	AL928553	AL928553 Mouse DNA
240	35	87.5	12311	6	A47690	A47690 Sequence 1	c 313	35	87.5	150292	9	AC112342	AC112342 Rattus no
241	35	87.5	12311	6	AR126227	AR126227 Sequence	c 314	35	87.5	150765	14	AC024326	AC024326 Homo sapi
242	35	87.5	12311	13	HCVPOLYPR	Z46258 Hog cholera	c 315	35	87.5	151230	14	AC150050	AC150050 Gallus ga
243	35	87.5	12854	1	TIPULSA	MS7692 Thermoanaer	c 316	35	87.5	151595	5	AL954191	AL954191 Zebrafish
244	35	87.5	12881	1	AE006091	AE006091 Pasteurel	c 317	35	87.5	152676	14	AC157610	AC157610 Mus muscu
245	35	87.5	17539	1	SAU73374	U73374 Staphylococ	c 318	35	87.5	153370	8	HSA119F19	AL132656 Human DNA
246	35	87.5	17950	15	AY372275	AY372275 Schizophy	c 319	35	87.5	155805	8	AL359265	AL359265 Human DNA
247	35	87.5	18131	1	SAU81973	U81973 Staphylococ	c 320	35	87.5	156165	9	AC158124	AC158124 Mus muscu
248	35	87.5	19236	13	AY340974	AY340974 Citrus tr	c 321	35	87.5	156542	14	AC150633	AC150633 Bos tauru
249	35	87.5	29546	2	AY246561	AY246561 Branchios	c 322	35	87.5	156551	8	AC106827	AC106827 Homo sapi
250	35	87.5	34223	6	AB044334	AB044334 Mus muscu	c 323	35	87.5	157289	14	AC021447	AC021447 Homo sapi
251	35	87.5	36221	6	AR669945	AR669945 Sequence	c 324	35	87.5	158139	5	BX120016	BX120016 Zebrafish
252	35	87.5	36532	8	AP001223	AP001223 Homo sapi	c 325	35	87.5	160920	8	AP000403	AP000403 Homo sapi
253	35	87.5	40987	2	AF391296	AF391296 Branchios	c 326	35	87.5	161276	8	AC012690	AC012690 Homo sapi
254	35	87.5	42133	15	YSCD9481	Y28373 Saccharomyc	c 327	35	87.5	161723	8	AC079790	AC079790 Homo sapi
255	35	87.5	43178	9	AC004405	AC004405 Mus muscu	c 328	35	87.5	161998	9	AC154780	AC154780 Mus muscu
256	35	87.5	43576	9	AC003994	AC003994 Mouse cos	c 329	35	87.5	162121	5	AL935311	AL935311 Zebrafish
257	35	87.5	48915	5	BX323554	BX323554 Zebrafish	c 330	35	87.5	162286	8	AL157893	AL157893 Human DNA
258	35	87.5	50752	14	AC100249	AC100249 Mus muscu	c 331	35	87.5	163132	8	AC025038	AC025038 Homo sapi
259	35	87.5	62988	8	AL731680	AL731680 Human DNA	c 332	35	87.5	164782	14	AC150052	AC150052 Gallus ga
260	35	87.5	74427	8	AL358796	AL358796 Human DNA	c 333	35	87.5	165268	14	AC149470	AC149470 Rhinoph
261	35	87.5	77860	15	ATT10C21	AL109787 Arabidops	c 334	35	87.5	165308	14	AC166030	AC166030 Sus scrof
262	35	87.5	77939	15	AB005246	AB005246 Arabidops	c 335	35	87.5	167278	14	AP001536	AP001536 Homo sapi
263	35	87.5	84570	8	AL359552	AL359552 Human DNA	c 336	35	87.5	168721	9	AC125085	AC125085 Mus muscu
264	35	87.5	84679	14	AC022571	AC022571 Homo sapi	c 337	35	87.5	169484	8	AC097478	AC097478 Homo sapi
265	35	87.5	92836	8	AL603908	AL603908 Human DNA	c 338	35	87.5	171161	9	AC105980	AC105980 Mus muscu
266	35	87.5	93033	14	AP000644	AP000644 Homo sapi	c 339	35	87.5	171402	9	AC140437	AC140437 Mus muscu
267	35	87.5	94884	14	AP007476	AP007476 Lotus cor	c 340	35	87.5	172759	8	AC025589	AC025589 Homo sapi
268	35	87.5	95153	8	AC125786	AC125786 Homo sapi	c 341	35	87.5	173257	9	AC154024	AC154024 Mus muscu
269	35	87.5	98179	8	AC092363	AC092363 Homo sapi	c 342	35	87.5	174897	14	BX957304	BX957304 Danio rer
270	35	87.5	99827	5	BX323581	BX323581 Zebrafish	c 343	35	87.5	175319	9	AC134837	AC134837 Mus muscu
271	35	87.5	101427	8	AL596269	AL596269 Human DNA	c 344	35	87.5	175644	8	AC092596	AC092596 Homo sapi
272	35	87.5	107438	8	AC005281	AC005281 Homo sapi	c 345	35	87.5	176432	14	AC114419	AC114419 Mus muscu
273	35	87.5	110000	1	CR628337_05	Continuation (6 of	c 346	35	87.5	176653	14	AP001857	AP001857 Homo sapi
274	35	87.5	110000	1	AE008692_12	Continuation (13 o	c 347	35	87.5	176736	5	AL929337	AL929337 Zebrafish
275	35	87.5	110000	1	AE007180_37	Continuation (38 o	c 348	35	87.5	176759	14	AC037464	AC037464 Homo sapi
276	35	87.5	110000	1	AP006841_33	Continuation (34 o	c 349	35	87.5	176910	8	AC112512	AC112512 Homo sapi
277	35	87.5	110000	1	AP008934_03	Continuation (4 of	c 350	35	87.5	177011	14	CT009715	CT009715 Mus muscu
278	35	87.5	110000	1	BA000017_01	Continuation (2 of	c 351	35	87.5	177038	14	AC128873	AC128873 Rattus no
279	35	87.5	110000	1	BA000018_01	Continuation (2 of	c 352	35	87.5	177433	8	AP001458	AP001458 Homo sapi
280	35	87.5	110000	1	BA000026_10	Continuation (11 o	c 353	35	87.5	177990	14	AC090306	AC090306 Homo sapi
281	35	87.5	110000	1	BA000033_01	Continuation (2 of	c 354	35	87.5	179652	8	AC022878	AC022878 Homo sapi
282	35	87.5	110000	1	BX571856_01	Continuation (2 of	c 355	35	87.5	180496	14	AC113875	AC113875 Rattus no
283	35	87.5	110000	1	EX571857_01	Continuation (2 of	c 356	35	87.5	181632	14	AC053508	AC053508 Homo sapi
284	35	87.5	110000	1	CP000009_18	Continuation (19 o	c 357	35	87.5	181878	9	AL683827	AL683827 Mouse DNA
285	35	87.5	110000	1	CP000009_19	Continuation (20 o	c 358	35	87.5	182079	14	OSJN00133	AL662943 Oryza sat
286	35	87.5	110000	1	CP000023_11	Continuation (12 o	c 359	35	87.5	182258	9	AC113969	AC113969 Mus muscu
287	35	87.5	110000	1	CP000024_11	Continuation (12 o	c 360	35	87.5	182835	14	AC069569	AC069569 Homo sapi
288	35	87.5	110000	1	CP000033_08	Continuation (9 of	c 361	35	87.5	182855	9	AC132086	AC132086 Mus muscu
289	35	87.5	110000	1	CP000046_01	Continuation (2 of	c 362	35	87.5	183666	8	AC009487	AC009487 Homo sapi
290	35	87.5	110000	1	CP000076_13	Continuation (14 o	c 363	35	87.5	185624	8	AC108162	AC108162 Homo sapi
291	35	87.5	110000	9	AE008684_0	AE008684 Mus muscu	c 364	35	87.5	185786	14	CR954228	CR954228 Danio rer
292	35	87.5	110000	9	AE008685_0	AE008685 Mus muscu	c 365	35	87.5	187677	14	AC147321	AC147321 Pan trogl
293	35	87.5	110000	9	AE01482_2	Continuation (3 of	c 366	35	87.5	187691	9	AC117220	AC117220 Mus muscu
294	35	87.5	110000	15	AP008218_192	Continuation (193	c 367	35	87.5	187741	14	AC087681	AC087681 Homo sapi
295	35	87.5	110000	15	AP008210_028	Continuation (29 o	c 368	35	87.5	189167	5	BX890568	BX890568 Zebrafish

[illegible]

515	34	85.0	14456	1	AE002339	AE002339 Chlamydia	588	34	85.0	110000	1	CP000090_01	Continuation (2 of
c 516	34	85.0	15699	8	AC003076	AC003076 Homo sapi	589	34	85.0	110000	2	AE003789_2	Continuation (3 of
517	34	85.0	18734	9	AY188393	AY188393 Cricetulu	590	34	85.0	110000	6	AR310754_11	Continuation (12 o
c 518	34	85.0	22589	1	AE002240	AE002240 Chlamydog	591	34	85.0	110000	6	AR607478_11	Continuation (12 o
519	34	85.0	24347	8	AL356964	AL356964 Human DNA	592	34	85.0	110000	14	AC108277_2	Continuation (3 of
c 520	34	85.0	24639	14	AC080061	AC080061 Homo sapi	c 593	34	85.0	110000	14	AC128823_1	Continuation (2 of
c 521	34	85.0	25025	2	CER04B5	Z70782 Caenorhabdi	c 594	34	85.0	110000	14	AC151276_1	Continuation (2 of
522	34	85.0	26875	8	AV044273	AV044273 Homo sapi	c 595	34	85.0	110000	14	AL672261_0	AL672261 Mus muscu
523	34	85.0	28432	6	CQ576998	CQ576998 Sequence	c 596	34	85.0	110000	14	CR933017_2	Continuation (3 of
524	34	85.0	30894	2	CEFS7A8	Z70781 Caenorhabdi	c 597	34	85.0	110000	15	AP008215_007	Continuation (8 of
c 525	34	85.0	37346	8	AC137777	AC137777 Homo sapi	598	34	85.0	110000	15	AP008216_164	Continuation (165
526	34	85.0	37474	8	HSJ647012	AL121922 Human DNA	599	34	85.0	110000	15	CR382130_27	Continuation (28 o
527	34	85.0	37972	14	AC137601	AC137601 Homo sapi	600	34	85.0	110000	15	AP008208_230	Continuation (231
c 528	34	85.0	38206	14	AC166900	AC166900 Mus muscu	601	34	85.0	110000	15	AP008209_173	Continuation (174
c 529	34	85.0	39407	14	AC137636	AC137636 Homo sapi	c 602	34	85.0	110000	15	AP008210_206	Continuation (207
530	34	85.0	42300	2	U80452	U80452 Caenorhabdi	c 603	34	85.0	110000	15	AP008210_235	Continuation (236
531	34	85.0	56179	15	AP004626	AP004626 Lotus cor	c 604	34	85.0	110000	15	AP008211_091	Continuation (92 o
c 532	34	85.0	62359	14	AC116167	AC116167 Homo sapi	c 605	34	85.0	110000	15	AP008211_092	Continuation (93 o
533	34	85.0	62623	8	AC080068	AC080068 Homo sapi	c 606	34	85.0	110000	15	AP008211_264	Continuation (265
c 534	34	85.0	63241	14	AC100715	AC100715 Mus muscu	c 607	34	85.0	110000	15	AP008211_278	Continuation (279
c 535	34	85.0	64168	14	AC079003	AC079003 Homo sapi	608	34	85.0	110000	15	AP008211_292	Continuation (292
536	34	85.0	64781	14	AC102025	AC102025 Homo sapi	c 609	34	85.0	110000	15	AP008212_046	Continuation (47 o
537	34	85.0	65311	8	AL162422	AL162422 Human DNA	c 610	34	85.0	111123	14	AC066604	AC066604 Homo sapi
538	34	85.0	66917	14	AC021268	AC021268 Homo sapi	611	34	85.0	111301	8	AC138331	AL338331 Homo sapi
539	34	85.0	73721	14	AC016531	AC016531 Homo sapi	612	34	85.0	113539	8	AL353681	AL353681 Human DNA
c 540	34	85.0	75547	8	AC004919	AC004919 Homo sapi	613	34	85.0	114428	8	AC092468	AC092468 Homo sapi
541	34	85.0	80646	14	AC013962	AC013962 Drosophil	614	34	85.0	115109	8	AC107216	AC107216 Homo sapi
c 542	34	85.0	82206	14	AC106541_4	Continuation (5 of	c 615	34	85.0	118533	8	AC008436	AC008436 Homo sapi
c 543	34	85.0	83373	15	AB017064	AB017064 Arabidops	c 616	34	85.0	118610	14	AC002344	AC002344 Homo sapi
544	34	85.0	85424	8	AC091179	AC091179 Homo sapi	617	34	85.0	118767	9	BX000344	BX000344 Mouse DNA
c 545	34	85.0	87000	8	AC097103	AC097103 Homo sapi	c 618	34	85.0	119126	15	AC149050	AC149050 Medicago
c 546	34	85.0	87118	8	AL355372	AL355372 Human DNA	619	34	85.0	120311	9	AC006945	AC006945 Mus muscu
547	34	85.0	90840	15	AB073158	AB073158 Arabidops	c 620	34	85.0	124192	5	BX323855	BX323855 Zebrafish
548	34	85.0	91039	14	AC165551	AC165551 Bos tauru	c 621	34	85.0	124324	14	AL158138	AL158138 Homo sapi
549	34	85.0	92394	14	AC018307	AC018307 Drosophil	c 622	34	85.0	124629	15	OSJN00062	AL606618 Oryza sat
c 550	34	85.0	93393	9	AL018483	AL018483 Sus scrof	c 623	34	85.0	126329	9	AL928544	AL928544 Mouse DNA
551	34	85.0	94890	9	AL591843	AL591843 Mouse DNA	624	34	85.0	127051	8	AL355538	AL355538 Human DNA
552	34	85.0	95976	8	AC068062	AC068062 Homo sapi	c 625	34	85.0	128953	14	AC155410	AC155410 Zea mays
c 553	34	85.0	97032	9	AL928869	AL928869 Mouse DNA	626	34	85.0	128960	8	HSJ919F19	AL109947 Human DNA
c 554	34	85.0	97155	9	AL845355	AL845355 Mouse DNA	c 627	34	85.0	129949	14	AC022940	AC022940 Homo sapi
555	34	85.0	97978	8	AL513285	AL513285 Human DNA	c 628	34	85.0	129957	8	HS415G2	Z83846 Human DNA s
c 556	34	85.0	98669	14	AC149823	AC149823 Zea mays	629	34	85.0	131070	8	AL353744	AL353744 Human DNA
557	34	85.0	99460	14	AC106346_3	Continuation (4 of	c 630	34	85.0	131181	14	CT010459	CT010459 Medicago
558	34	85.0	100525	15	AP004079	AP004079 Oryza sat	c 631	34	85.0	131227	14	AL136098	AL136098 Homo sapi
559	34	85.0	100800	15	AC105260	AC105260 Oryza sat	c 632	34	85.0	131894	8	AC108751	AC108751 Homo sapi
c 560	34	85.0	100848	8	AV842481	AV842481 Homo sapi	c 633	34	85.0	132977	14	RN510D20	AL603723 Rattus no
561	34	85.0	101652	8	AC079754	AC079754 Homo sapi	c 634	34	85.0	134257	5	BX006321	BX006321 Zebrafish
c 562	34	85.0	102602	14	AC149581	AC149581 Medicago	635	34	85.0	135983	8	HSJ595K12	AL096829 Human DNA
563	34	85.0	105836	8	AC068780	AC068780 Homo sapi	636	34	85.0	137220	15	AC135598	AL135598 Oryza sat
c 564	34	85.0	106886	8	AP001462	AP001462 Homo sapi	c 637	34	85.0	138615	15	AP004023	AP004023 Oryza sat
565	34	85.0	107066	14	AC138264	AC138264 Homo sapi	638	34	85.0	138663	9	AL713874	AL713874 Mouse DNA
c 566	34	85.0	107076	8	AL159997	AL159997 Human DNA	c 639	34	85.0	138675	8	HS1057B20	AL109823 Human DNA
567	34	85.0	107252	14	CT009722	CT009722 Danio rer	c 640	34	85.0	139298	15	OSJN00012	AL606441 Oryza sat
568	34	85.0	110000	1	AE006470_02	Continuation (3 of	c 641	34	85.0	139722	14	AC154137	AC154137 Tribolium
c 569	34	85.0	110000	1	CR522870_07	Continuation (8 of	642	34	85.0	140425	8	AC005479	AC005479 Homo sapi
c 570	34	85.0	110000	1	CR522870_27	Continuation (28 o	c 643	34	85.0	140432	14	AC166125	AC166125 Mus muscu
c 571	34	85.0	110000	1	CR555306_16	Continuation (17 o	c 644	34	85.0	141019	14	AC162144	AC162144 Loxodonta
572	34	85.0	110000	1	AE014295_03	Continuation (4 of	c 645	34	85.0	142854	15	AC136216	AC136216 Oryza sat
c 573	34	85.0	110000	1	AE014295_09	Continuation (10 o	646	34	85.0	143406	9	AC092857	AC092857 Rattus no
c 574	34	85.0	110000	1	AE016828_12	Continuation (13 o	c 647	34	85.0	147704	15	AC097112	AC097112 Oryza sat
c 575	34	85.0	110000	1	AE017199_1	Continuation (2 of	648	34	85.0	148416	8	AC002089	AC002089 Homo sapi
c 576	34	85.0	110000	1	AP006627_39	Continuation (40 o	c 649	34	85.0	148673	8	AC092629	AC092629 Homo sapi
577	34	85.0	110000	1	AP006716_05	Continuation (6 of	c 650	34	85.0	149202	14	AC073467	AC073467 Homo sapi
c 578	34	85.0	110000	1	AP006716_23	Continuation (24 o	651	34	85.0	149339	9	AC125112	AC125112 Mus muscu
579	34	85.0	110000	1	AP008226_16	Continuation (17 o	c 652	34	85.0	149489	14	AC149451	AC149451 Papio annu
580	34	85.0	110000	1	AP008934_00	AP008934 Staphyloc	653	34	85.0	149978	15	AC068951	AC068951 Oryza sat
581	34	85.0	110000	1	AP008934_03	Continuation (4 of	c 654	34	85.0	150604	15	AP005787	AP005787 Oryza sat
582	34	85.0	110000	1	BA000008_11	Continuation (12 o	c 655	34	85.0	150726	8	AC092991	AC092991 Homo sapi
583	34	85.0	110000	1	BA000015_03	Continuation (4 of	c 656	34	85.0	151030	9	AL805936	AL805936 Mouse DNA
c 584	34	85.0	110000	1	BA000016_12	Continuation (13 o	657	34	85.0	151440	5	BX510933	BX510933 Zebrafish
c 585	34	85.0	110000	1	BA000022_27	Continuation (28 o	c 586	34	85.0	151571	14	AC157269	AC157269 Rhinoph
586	34	85.0	110000	1	CP000023_09	Continuation (10 o	c 587	34	85.0	151577	9	AC133156	AC133156 Mus muscu
587	34	85.0	110000	1	CP000024_09	Continuation (10 o	c 660	34	85.0	151744	14	AC150201	AC150201 Papio annu

661	34	85.0	151766	9	AC154006	AC154006 Mus muscu	c 734	34	85.0	174046	14	AC159067	AC159067 Bos tauru
662	34	85.0	152228	8	AL135902	AL135902 Human DNA	c 735	34	85.0	174500	15	AP004738	AP004738 Oryza sat
663	34	85.0	152289	8	AC087430	AC087430 Homo sapi	c 736	34	85.0	174560	9	AC101944	AC101944 Mus muscu
664	34	85.0	152635	14	AC011144	AC011144 Homo sapi	c 737	34	85.0	175345	14	AC022918	AC022918 Homo sapi
665	34	85.0	153787	14	AC023037	AC023037 Homo sapi	c 738	34	85.0	175370	14	AC145861	AC145861 Pan trogl
666	34	85.0	154123	14	AC019562	AC019562 Drosophil	c 739	34	85.0	175604	9	AC131756	AC131756 Mus muscu
667	34	85.0	154359	9	AC110501	AC110501 Mus muscu	c 740	34	85.0	175641	9	AC147556	AC147556 Mus muscu
668	34	85.0	154439	9	AC016646	AC016646 Homo sapi	c 741	34	85.0	175757	8	AL356534	AL356534 Human DNA
669	34	85.0	154495	14	AC161739	AC161739 Dasyypus n	c 742	34	85.0	175762	5	AL772332	AL772332 zebrafish
670	34	85.0	154369	8	AL160058	AL160058 Human DNA	c 743	34	85.0	175837	14	AC022045	AC022045 Homo sapi
671	34	85.0	155646	8	AC146044	AC146044 Pan trogl	c 744	34	85.0	175944	14	AC026878	AC026878 Mus muscu
672	34	85.0	156248	9	AC125443	AC125443 Mus muscu	c 745	34	85.0	176521	2	AC097725	AC097725 Drosophil
673	34	85.0	156426	14	AC152022	AC152022 Papio anu	c 746	34	85.0	176844	14	AC120333	AC120333 Rattus no
674	34	85.0	157106	15	AC124143	AC124143 Oryza sat	c 747	34	85.0	176982	8	AC092646	AC092646 Homo sapi
675	34	85.0	157171	14	AC155454	AC155454 Zea mays	c 748	34	85.0	177308	8	AC009560	AC009560 Homo sapi
676	34	85.0	157106	8	AL603902	AL603902 Human DNA	c 749	34	85.0	177661	9	AC119237	AC119237 Mus muscu
677	34	85.0	158024	8	AL603902	AL603902 Human DNA	c 750	34	85.0	177784	8	AC027465	AC027465 Homo sapi
678	34	85.0	158508	8	AC107622	AC107622 Homo sapi	c 751	34	85.0	178445	5	CR376771	CR376771 Zebrafish
679	34	85.0	158745	9	AC131801	AC131801 Mus muscu	c 752	34	85.0	178445	5	AC096225	AC096225 Homo sapi
680	34	85.0	159071	15	AP005591	AP005591 Oryza sat	c 753	34	85.0	179357	14	AL929338	AL929338 zebrafish
681	34	85.0	159285	8	AL354816	AL354816 Human DNA	c 754	34	85.0	179399	5	AL929338	AL929338 zebrafish
682	34	85.0	159294	14	AC141974	AC141974 Rattus no	c 755	34	85.0	180162	2	AC007582	AC007582 Drosophil
683	34	85.0	159479	14	AC126342	AC126342 Homo sapi	c 756	34	85.0	180931	9	AC133193	AC133193 Mus muscu
684	34	85.0	161218	5	AL772148	AL772148 Zebrafish	c 757	34	85.0	181070	14	AC142187	AC142187 Rattus no
685	34	85.0	161268	14	AL158700	AL158700 Oryza sat	c 758	34	85.0	181106	9	AL929011	AL929011 Mouse DNA
686	34	85.0	161514	9	BX004793	BX004793 Mouse DNA	c 759	34	85.0	181148	8	AC119427	AC119427 Homo sapi
687	34	85.0	161771	5	BX072576	BX072576 Zebrafish	c 760	34	85.0	181348	14	AC123917	AC123917 Mus muscu
688	34	85.0	162197	9	AC139333	AC139333 Mus muscu	c 761	34	85.0	181942	8	AC079227	AC079227 Homo sapi
689	34	85.0	162463	9	AC102525	AC102525 Mus muscu	c 762	34	85.0	182555	8	AC068531	AC068531 Homo sapi
690	34	85.0	162770	9	AC121904	AC121904 Mus muscu	c 763	34	85.0	182627	14	AC024262	AC024262 Homo sapi
691	34	85.0	162863	14	AC161115	AC161115 Mus muscu	c 764	34	85.0	182776	8	AC091078	AC091078 Homo sapi
692	34	85.0	163095	14	AC149251	AC149251 Papio anu	c 765	34	85.0	182868	14	AC121360	AC121360 Oryza sat
693	34	85.0	163367	9	AC104518	AC104518 Mus muscu	c 766	34	85.0	183320	14	AC163769	AC163769 Pan trogl
694	34	85.0	163503	9	AC119253	AC119253 Mus muscu	c 767	34	85.0	183359	14	AC161574	AC161574 Mus muscu
695	34	85.0	163688	9	AC105979	AC105979 Mus muscu	c 768	34	85.0	185365	9	AC069559	AC069559 Mus muscu
696	34	85.0	163806	8	AC113349	AC113349 Homo sapi	c 769	34	85.0	186045	9	AC123830	AC123830 Mus muscu
697	34	85.0	164000	8	AC121343	AC121343 Homo sapi	c 770	34	85.0	186151	9	AL671042	AL671042 Mouse DNA
698	34	85.0	164049	14	AC151086	AC151086 Bos tauru	c 771	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
699	34	85.0	164356	2	AC012166	AC012166 Drosophil	c 772	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
700	34	85.0	164925	15	AC121362	AC121362 Oryza sat	c 773	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
701	34	85.0	166197	14	AC073552	AC073552 Homo sapi	c 774	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
702	34	85.0	166232	9	AC122400	AC122400 Mus muscu	c 775	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
703	34	85.0	166406	9	AC124005	AC124005 Mus muscu	c 776	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
704	34	85.0	166720	8	AC112180	AC112180 Homo sapi	c 777	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
705	34	85.0	166857	14	AC098589	AC098589 Homo sapi	c 778	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
706	34	85.0	166960	14	AC118180	AC118180 Rattus no	c 779	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
707	34	85.0	167179	8	AC026182	AC026182 Homo sapi	c 780	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
708	34	85.0	167273	9	AC083894	AC083894 Mus muscu	c 781	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
709	34	85.0	167315	14	AC152375	AC152375 Dasyypus n	c 782	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
710	34	85.0	167486	8	AC098615	AC098615 Homo sapi	c 783	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
711	34	85.0	168140	14	AC013379	AC013379 Homo sapi	c 784	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
712	34	85.0	168805	14	AC027525	AC027525 Homo sapi	c 785	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
713	34	85.0	168872	14	AC016296	AC016296 Homo sapi	c 786	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
714	34	85.0	168900	14	AC026969	AC026969 Homo sapi	c 787	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
715	34	85.0	169480	14	AC068074	AC068074 Homo sapi	c 788	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
716	34	85.0	169614	9	AC123817	AC123817 Mus muscu	c 789	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
717	34	85.0	169659	8	AC024933	AC024933 Homo sapi	c 790	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
718	34	85.0	169661	9	AC132344	AC132344 Mus muscu	c 791	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
719	34	85.0	170056	9	AC100727	AC100727 Mus muscu	c 792	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
720	34	85.0	171114	8	AC092757	AC092757 Homo sapi	c 793	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
721	34	85.0	171559	14	BX511031	BX511031 Danio rer	c 794	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
722	34	85.0	171612	14	AC144193	AC144193 Macaca mu	c 795	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
723	34	85.0	171817	14	AC027681	AC027681 Mus muscu	c 796	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
724	34	85.0	171982	14	AC120916	AC120916 Rattus no	c 797	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
725	34	85.0	172914	8	AC015604	AC015604 Homo sapi	c 798	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
726	34	85.0	172959	5	BX119985	BX119985 Zebrafish	c 799	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
727	34	85.0	173238	14	AC006489	AC006489 Drosophil	c 800	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
728	34	85.0	173467	5	BX571699	BX571699 Zebrafish	c 801	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
729	34	85.0	173691	14	AC087462	AC087462 Homo sapi	c 802	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
730	34	85.0	173740	14	AC020964	AC020964 Mus muscu	c 803	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
731	34	85.0	173854	8	AC090043	AC090043 Homo sapi	c 804	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
732	34	85.0	173872	8	AP000428	AP000428 Homo sapi	c 805	34	85.0	186200	9	AL671042	AL671042 Mouse DNA
733	34	85.0	173979	14	AC149625	AC149625 Papio anu	c 806	34	85.0	186200	9	AL671042	AL671042 Mouse DNA

807	34	85.0	195691	14	AC159905	AC159905 Pan trogl	880	34	85.0	223397	14	BX001052	BX001052 Mus muscu
808	34	85.0	195710	9	AC161117	AC161117 Mus muscu	891	34	85.0	223575	14	CT009747	CT009747 Mus muscu
809	34	85.0	195895	14	AC117033	AC117033 Rattus no	c 882	34	85.0	223670	1	CR555308	CR555308 Azoarcus
810	34	85.0	196050	1	AL646058	AL646058 Raletonia	c 883	34	85.0	223711	14	AC126979	AC126979 Rattus no
811	34	85.0	196398	9	AL604027	AL604027 Mouse DNA	884	34	85.0	224009	14	AC131058	AC131058 Mus muscu
812	34	85.0	196591	14	AC091082	AC091082 Homo sapi	885	34	85.0	224266	14	AC127197	AC127197 Rattus no
813	34	85.0	196919	14	AC118165	AC118165 Rattus no	886	34	85.0	224935	9	AC150275	AC150275 Mus muscu
814	34	85.0	197032	14	AC159092	AC159092 Mus muscu	c 887	34	85.0	225002	9	AL824707	AL824707 Mouse DNA
815	34	85.0	197371	14	AL355524	AL355524 Homo sapi	c 888	34	85.0	225260	14	AC095002	AC095002 Rattus no
816	34	85.0	197457	14	AC155133	AC155133 Bos tauru	c 889	34	85.0	225558	14	AC103066	AC103066 Rattus no
817	34	85.0	198478	14	AC154347	AC154347 Mus muscu	c 890	34	85.0	225763	14	AC115241	AC115241 Rattus no
818	34	85.0	198543	14	AC163672	AC163672 Mus muscu	c 891	34	85.0	226177	14	AC123959	AC123959 Rattus no
819	34	85.0	198548	14	AC161618	AC161618 Pan trogl	c 892	34	85.0	226218	14	AC113752	AC113752 Rattus no
820	34	85.0	198917	8	AL356157	AL356157 Human DNA	c 893	34	85.0	226280	14	AC099657	AC099657 Rattus no
821	34	85.0	199086	14	AC132250	AC132250 Mus muscu	894	34	85.0	226345	8	AC005406	AC005406 Homo sapi
822	34	85.0	199446	9	AC135069	AC135069 Mus muscu	895	34	85.0	226786	9	AL731896	AL731896 Mouse DNA
823	34	85.0	200069	14	AC127190	AC127190 Rattus no	c 896	34	85.0	227090	14	AC099391	AC099391 Rattus no
824	34	85.0	200799	15	AF137379	AF137379 Nephrosel	c 897	34	85.0	227958	14	AC109418	AC109418 Rattus no
825	34	85.0	202269	9	AC102033	AC102033 Mus muscu	c 898	34	85.0	228008	2	AE003466	AE003466 Drosophil
826	34	85.0	202402	9	AV189120	AV189120 Mus muscu	899	34	85.0	228327	5	AL844514	AL844514 Zebrafish
827	34	85.0	202544	9	AC113527	AC113527 Mus muscu	c 900	34	85.0	228624	5	BX088524	BX088524 Zebrafish
828	34	85.0	202606	14	AC164562	AC164562 Mus muscu	901	34	85.0	228745	14	AC128310	AC128310 Rattus no
829	34	85.0	202884	5	BX548019	BX548019 Zebrafish	c 902	34	85.0	229240	14	AL590291	AL590291 Homo sapi
830	34	85.0	202887	9	AL928789	AL928789 Mouse DNA	c 903	34	85.0	229283	14	AC095445	AC095445 Rattus no
831	34	85.0	203300	8	AC000134	AC000134 Homo sapi	c 904	34	85.0	229321	14	AC094185	AC094185 Rattus no
832	34	85.0	203685	9	AC122339	AC122339 Mus muscu	905	34	85.0	229478	14	AC151871	AC151871 Lemur cat
833	34	85.0	203753	8	AC104343	AC104343 Homo sapi	906	34	85.0	229616	14	AC156070	AC156070 Bos tauru
834	34	85.0	204032	1	AL646070	AL646070 Raletonia	c 907	34	85.0	229728	14	BX936294	BX936294 Mus muscu
835	34	85.0	204050	14	AC148411	AC148411 Oroleum	c 908	34	85.0	230148	14	AC094915	AC094915 Rattus no
836	34	85.0	204331	14	AC160670	AC160670 Bos tauru	c 909	34	85.0	230562	14	AC160250	AC160250 Bos tauru
837	34	85.0	204590	14	AC102754	AC102754 Mus muscu	c 910	34	85.0	231544	14	AC112525	AC112525 Rattus no
838	34	85.0	205085	15	ATPCN4	297339 Arabidopsis	c 911	34	85.0	231864	14	AC098112	AC098112 Rattus no
839	34	85.0	205166	14	AC148169	AC148169 Zea mays	c 912	34	85.0	232191	14	AC096110	AC096110 Rattus no
840	34	85.0	205503	14	AC141489	AC141489 Rattus no	c 913	34	85.0	232310	14	AC109074	AC109074 Rattus no
841	34	85.0	205819	9	AC131654	AC131654 Mus muscu	c 914	34	85.0	232544	14	AC129393	AC129393 Rattus no
842	34	85.0	206059	14	AC114986	AC114986 Mus muscu	c 915	34	85.0	232779	14	AC093933	AC093933 Rattus no
843	34	85.0	206380	9	AC113538	AC113538 Mus muscu	c 916	34	85.0	234008	14	CR392350	CR392350 Danio rer
844	34	85.0	207420	14	AC078884	AC078884 Mus muscu	917	34	85.0	234766	14	AC111716	AC111716 Rattus no
845	34	85.0	207441	14	AC140719	AC140719 Homo sapi	918	34	85.0	234882	14	AC127871	AC127871 Rattus no
846	34	85.0	207842	8	AC101076	AC101076 Homo sapi	919	34	85.0	234961	14	AC109380	AC109380 Rattus no
847	34	85.0	208483	9	AC107757	AC107757 Mus muscu	c 920	34	85.0	235381	14	AC103031	AC103031 Rattus no
848	34	85.0	208895	8	AC159715	AC159715 Mus muscu	921	34	85.0	235445	5	BX950179	BX950179 Zebrafish
849	34	85.0	208912	8	AC147022	AC147022 Pan trogl	922	34	85.0	236356	14	AC128343	AC128343 Rattus no
850	34	85.0	209764	9	AL645468	AL645468 Mouse DNA	923	34	85.0	236676	14	AC115491	AC115491 Rattus no
851	34	85.0	209836	9	BX284624	BX284624 Mouse DNA	c 924	34	85.0	236846	14	AC095590	AC095590 Rattus no
852	34	85.0	210532	8	AC008558	AC008558 Homo sapi	925	34	85.0	237178	14	AC136754	AC136754 Mus muscu
853	34	85.0	210571	14	CR854953	CR854953 Danio rer	c 926	34	85.0	237247	14	AC094990	AC094990 Rattus no
854	34	85.0	211085	14	AC096859	AC096859 Rattus no	c 927	34	85.0	237581	14	AC131646	AC131646 Rattus no
855	34	85.0	211163	14	AC141508	AC141508 Rattus no	928	34	85.0	237590	9	AC110240	AC110240 Mus muscu
856	34	85.0	211540	14	AC117161	AC117161 Rattus no	929	34	85.0	237737	9	AC123681	AC123681 Mus muscu
857	34	85.0	211682	5	AL772198	AL772198 Zebrafish	930	34	85.0	237754	9	AC125576	AC125576 Rattus no
858	34	85.0	211856	14	CT009572	CT009572 Mus muscu	c 931	34	85.0	237788	14	AC118498	AC118498 Rattus no
859	34	85.0	212435	5	BX908748	BX908748 Zebrafish	c 932	34	85.0	237881	14	AC105704	AC105704 Rattus no
860	34	85.0	213534	14	AL390792	AL390792 Homo sapi	c 933	34	85.0	238907	14	AC128162	AC128162 Rattus no
861	34	85.0	214192	14	AL354667	AL354667 Homo sapi	934	34	85.0	239047	14	AC163025	AC163025 Mus muscu
862	34	85.0	214567	9	AC105305	AC105305 Mus muscu	935	34	85.0	239600	9	AC139128	AC139128 Mus muscu
863	34	85.0	215225	8	AP001972	AP001972 Homo sapi	c 936	34	85.0	240344	14	AC127886	AC127886 Rattus no
864	34	85.0	215792	14	AC136190	AC136190 Rattus no	937	34	85.0	241045	14	AC161461	AC161461 Mus muscu
865	34	85.0	216048	9	AC141471	AC141471 Mus muscu	938	34	85.0	241518	14	AC098012	AC098012 Rattus no
866	34	85.0	216380	9	AC113113	AC113113 Mus muscu	c 939	34	85.0	241874	14	AC094162	AC094162 Rattus no
867	34	85.0	216932	9	AC090656	AC090656 Mus muscu	c 940	34	85.0	242954	9	AC142245	AC142245 Mus muscu
868	34	85.0	218926	14	AC146399	AC146399 Pan trogl	c 941	34	85.0	242974	14	AC103570	AC103570 Rattus no
869	34	85.0	219329	14	AC094590	AC094590 Rattus no	c 942	34	85.0	243078	14	AC126596	AC126596 Rattus no
870	34	85.0	219531	14	AC098098	AC098098 Rattus no	c 943	34	85.0	244136	14	AC126491	AC126491 Rattus no
871	34	85.0	219725	14	AC106324	AC106324 Rattus no	944	34	85.0	244301	14	AC160772	AC160772 Rattus no
872	34	85.0	219936	9	AC113474	AC113474 Mus muscu	945	34	85.0	244928	14	AC097177	AC097177 Rattus no
873	34	85.0	220236	14	AC110699	AC110699 Rattus no	c 946	34	85.0	245125	14	AC133759	AC133759 Rattus no
874	34	85.0	220314	14	AC097188	AC097188 Rattus no	947	34	85.0	245191	14	AC128815	AC128815 Rattus no
875	34	85.0	221454	13	AV186194	AV186194 Rhesus cy	948	34	85.0	245424	14	AC112822	AC112822 Rattus no
876	34	85.0	222186	9	AL611963	AL611963 Mouse DNA	949	34	85.0	246257	14	AC128192	AC128192 Rattus no
877	34	85.0	222392	9	AC116739	AC116739 Mus muscu	c 950	34	85.0	246387	14	AC109737	AC109737 Rattus no
878	34	85.0	222697	14	AC094121	AC094121 Rattus no	951	34	85.0	246966	14	AC119440	AC119440 Rattus no
879	34	85.0	223110	9	AC123752	AC123752 Mus muscu	c 952	34	85.0	247831	14	AC098072	AC098072 Rattus no

C 953 AC132729 Rattus no
C 954 AC120659 Rattus no
C 955 AC119126 Rattus no
C 956 AC106533 Rattus no
C 957 AC152758 Bos tauru
C 958 AC130036 Rattus no
C 959 CR388422 Denio rer
C 960 AC129272 Rattus no
C 961 AC093972 Rattus no
C 962 AC104053 Rattus no
C 963 AC126650 Rattus no
C 964 AC114706 Rattus no
C 965 AC156671 Bos tauru
C 966 BX004878 Zebrafish
C 967 AC073798 Mus muscu
C 968 AC098609 Rattus no
C 969 AC118627 Mus muscu
C 970 AC128269 Rattus no
C 971 AC123984 Mus muscu
C 972 AC125574 Rattus no
C 973 AL662812 Mouse DNA
C 974 AC163511 Bos tauru
C 975 AC112092 Rattus no
C 976 AC110346 Rattus no
C 977 AC126820 Rattus no
C 978 AC129791 Rattus no
C 979 AC111510 Rattus no
C 980 AC123080 Rattus no
C 981 AC116237 Rattus no
C 982 AC087226 Mus muscu
C 983 AC129399 Rattus no
C 984 AC133407 Rattus no
C 985 AC127885 Rattus no
C 986 AR408762 Sequence
C 987 AX067466 Sequence
C 988 AC107165 Rattus no
C 989 AC115502 Rattus no
C 990 AC094667 Rattus no
C 991 AC095925 Rattus no
C 992 AC128500 Rattus no
C 993 AC107003 Rattus no
C 994 AB017305 Thermus t
C 995 AC094114 Rattus no
C 996 AC105864 Rattus no
C 997 AC095358 Rattus no
C 998 AC131607 Rattus no
C 999 AC098526 Rattus no
1000 AC156662 Bos tauru

ALIGNMENTS

RESULT 1
LOCUS CQ687716 290 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 32642 from Patent WO02070737.
ACCESSION CQ687716
VERSION CQ687716.1 GI:42218962
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 32642 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source Location/Qualifiers
1..290
/organism="Homo sapiens"
/mol_type="unassigned DNA"

/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 12 Length: 290
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-10 (1-9) x CQ687716 (1-290)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 53 TTCTCGGGTATTGTTTACCCCTGATA 79
RESULT 2
LOCUS CQ920916 475 bp DNA linear PAT 23-NOV-2004
DEFINITION Sequence 2116 from Patent WO2004097052.
ACCESSION CQ920916
VERSION CQ920916.1 GI:56210857
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L., Strahs,A., Immerman,F. and Dörner,A.J.
TITLE Methods for prognosis and treatment of solid tumors
JOURNAL Patent: WO 2004097052-A 2116 11-NOV-2004;
Wyeth (US); Burczynski, Michael E. (US)
FEATURES
source Location/Qualifiers
1..475
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 19.1 Length: 475
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-10 (1-9) x CQ920916 (1-475)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 351 TTCTCGGGTATTGTTTACCCCTGATA 377
RESULT 3
LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 57 from Patent WO2059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Mack,D.H., Gish,K.C. and Afar,D.
TITLE Methods of diagnosis of breast cancer, compositions and methods of screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;

```

FEATURES             source
    source
    Location/Qualifiers
    1..927
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      Length: 927
Pred. No.:            35.7   Matches: 9
Score:                40.00   Conservatives: 0
Percent Similarity:   100.0%   Mismatches: 0
Best Local Similarity: 100.0%   Indels: 0
Query Match:         100.0%   Gaps: 0
DB:                   6

US-10-774-176-10 (1-9) x AX829164 (1-927)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 730 TTCTCGGTATTGTTTAGCCCTGATA 756

RESULT 4
AX467373
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
    Felis sp.
    Felis sp.
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
    Felinae; Felis.
REFERENCE
1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)

FEATURES             source
    source
    Location/Qualifiers
    1..1260
    /organism="Felis sp."
    /mol_type="unassigned DNA"
    /db_xref="taxon:9687"

ORIGIN
Alignment Scores:      Length: 1260
Pred. No.:            47.5   Matches: 9
Score:                40.00   Conservatives: 0
Percent Similarity:   100.0%   Mismatches: 0
Best Local Similarity: 100.0%   Indels: 0
Query Match:         100.0%   Gaps: 0
DB:                   6

US-10-774-176-10 (1-9) x AX467373 (1-1260)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1069 TTCTAGGTATTGTTTAGCCCTGATA 1095

RESULT 5
AX821533
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM
    Felis catus (cat)
    Felis catus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
    Felinae; Felis.
REFERENCE
1

AUTHORS Carroll, M., Kingman, S.M. and Redchenko, I.M.
TITLE MHC class I peptide epitopes from the human St4 tumor-associated antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)

FEATURES             source
    source
    Location/Qualifiers
    1..1260
    /organism="Felis catus"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9685"

ORIGIN
Alignment Scores:      Length: 1260
Pred. No.:            47.5   Matches: 9
Score:                40.00   Conservatives: 0
Percent Similarity:   100.0%   Mismatches: 0
Best Local Similarity: 100.0%   Indels: 0
Query Match:         100.0%   Gaps: 0
DB:                   6

US-10-774-176-10 (1-9) x AX821533 (1-1260)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1069 TTCTAGGTATTGTTTAGCCCTGATA 1095

RESULT 6
AX821548
LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE
ORGANISM
    Felis catus (cat)
    Felis catus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
    Felinae; Felis.
REFERENCE
1
AUTHORS Carroll, M.O., Harrop, R.O. and Kingsman, S.O.
TITLE MHC class II peptide epitope of St4 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)

FEATURES             source
    source
    Location/Qualifiers
    1..1260
    /organism="Felis catus"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9685"

ORIGIN
Alignment Scores:      Length: 1260
Pred. No.:            47.5   Matches: 9
Score:                40.00   Conservatives: 0
Percent Similarity:   100.0%   Mismatches: 0
Best Local Similarity: 100.0%   Indels: 0
Query Match:         100.0%   Gaps: 0
DB:                   6

US-10-774-176-10 (1-9) x AX821548 (1-1260)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 1069 TTCTAGGTATTGTTTAGCCCTGATA 1095

RESULT 7
BD249731
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 (bases 1 to 1263)
AUTHORS
Carroll, M.W. and Myers, K.A.

TITLE
Polypeptide
JOURNAL
Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD

COMMENT
OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002 JP 2000582415
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K14/065,
C07K19/00,
PC C12N15/00,
CC Polypeptide
FH Key Location/Qualifiers
FT source 1. 1263
/organism="Homo sapiens (human)".

FEATURES
source Location/Qualifiers
1. 1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 47.6 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x BD249731 (1-1263)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 1072 TTCTCGGTATTGTTTGGCCCTGATA 1098

RESULT 8
AX025011
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1
AUTHORS Carroll, M.W. and Myers, K.A.

TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)

FEATURES
source Location/Qualifiers
1. 1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 47.6 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX025011 (1-1263)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 1072 TTCTCGGTATTGTTTGGCCCTGATA 1098

RESULT 9
AX149553
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1
AUTHORS Kingsman, A.O., Kingsman, S.M., Bebbington, C.R., Carroll, M.W.,
Ellard, F.M. and Myers, K.A.

TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source Location/Qualifiers
1. 1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="574"

ORIGIN

Alignment Scores:
Pred. No.: 47.6 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX149553 (1-1263)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 1072 TTCTCGGTATTGTTTGGCCCTGATA 1098

RESULT 10
AX316086
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1
AUTHORS Carroll, M.W. and Myers, K.A.

TITLE St4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source Location/Qualifiers
1. 1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 47.6 Length: 1263


```

Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX316086 (1-1263)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1072 TTCTGGGTATTGTTTAGCCCTGATA 1098

RESULT 11
LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS
SOURCE
ORGANISM
Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
location/Qualifiers
1..1263
/mol_type="Canis sp."
/db_xref="taxon:9616"

ORIGIN
Alignment Scores:
Pred. No.: 47.6 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX467371 (1-1263)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1072 TTCTAGGTATTGTTTAGCCCTGATA 1098

RESULT 12
BD249732 1281 bp DNA linear PAT 17-JUL-2003
LOCUS BD249732
DEFINITION Polypeptide.
ACCESSION BD249732
VERSION BD249732.1 GI:33059502
KEYWORDS JP 2002530060-A/2.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1281)
Carroll, M.W. and Myers, K.A.
AUTHORS
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 2 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT
OS Mus musculus (mouse)
FN JP 2002530060-A/2
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4

Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX316086 (1-1263)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1072 TTCTGGGTATTGTTTAGCCCTGATA 1098

RESULT 11
LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS
SOURCE
ORGANISM
Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
location/Qualifiers
1..1263
/mol_type="Canis sp."
/db_xref="taxon:9616"

ORIGIN
Alignment Scores:
Pred. No.: 47.6 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX467371 (1-1263)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1072 TTCTAGGTATTGTTTAGCCCTGATA 1098

RESULT 12
BD249732 1281 bp DNA linear PAT 17-JUL-2003
LOCUS BD249732
DEFINITION Polypeptide.
ACCESSION BD249732
VERSION BD249732.1 GI:33059502
KEYWORDS JP 2002530060-A/2.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1281)
Carroll, M.W. and Myers, K.A.
AUTHORS
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 2 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT
OS Mus musculus (mouse)
FN JP 2002530060-A/2
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4

Score: 48.3 Length: 1281
Percent Similarity: 100.0% Matches: 9
Best Local Similarity: 100.0% Conservative: 0
Query Match: 100.0% Mismatches: 0
DB: 6 Indels: 0
Gaps: 0

US-10-774-176-10 (1-9) x AX025012 (1-1281)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1090 TTCTAGGTATTGTTTAGCTCTGATA 1116

RESULT 13
AX025012 1281 bp DNA linear PAT 15-SEP-2000
LOCUS AX025012
DEFINITION Sequence 2 from Patent WO029428.
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1
Carroll, M.W. and Myers, K.A.
AUTHORS
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 2 25-MAY-2000;
CARROLL, MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)

FEATURES
source
location/Qualifiers
1..1281
/mol_type="Mus musculus"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 48.3 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX025012 (1-1281)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1090 TTCTAGGTATTGTTTAGCTCTGATA 1116

RESULT 14
AX316087 1281 bp DNA linear PAT 14-DEC-2001
LOCUS AX316087
DEFINITION Sequence 2 from Patent EP1160323.
ACCESSION AX316087

```

```

PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K7/06, C07K14/065,
PC C07K19/00,
PC C12N15/00
CC Polypeptide
FH Key
FT source 1..1281
Location/Qualifiers
/organism="Mus musculus (mouse)".

FEATURES
source
location/Qualifiers
1..1281
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 48.3 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x BD249732 (1-1281)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1090 TTCTAGGTATTGTTTAGCTCTGATA 1116

RESULT 13
AX025012 1281 bp DNA linear PAT 15-SEP-2000
LOCUS AX025012
DEFINITION Sequence 2 from Patent WO029428.
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1
Carroll, M.W. and Myers, K.A.
AUTHORS
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 2 25-MAY-2000;
CARROLL, MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)

FEATURES
source
location/Qualifiers
1..1281
/mol_type="Mus musculus"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 48.3 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AX025012 (1-1281)

Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1090 TTCTAGGTATTGTTTAGCTCTGATA 1116

RESULT 14
AX316087 1281 bp DNA linear PAT 14-DEC-2001
LOCUS AX316087
DEFINITION Sequence 2 from Patent EP1160323.
ACCESSION AX316087

```

```

VERSION      AX316087.1  GI:17899279
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
KEYWORDS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
             Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE    1
AUTHORS      Carroll, M.W. and Myers, K.A.
TITLE        5T4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL      Patent: EP 1160323-A 2 05-DEC-2001;
             Oxford Biomedica (UK) Limited (GB)
FEATURES     source
             1..1281
               /organism="Mus musculus"
               /mol_type="unassigned DNA"
               /db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.:      48.3      Length:      1281
Score:          40.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             6        Gaps:       0

US-10-774-176-10 (1-9) x AX316087 (1-1281)
Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
Db 1090 TTCTAGGTATTGTTTGTAGCTCGATA 1116

RESULT 15
CQ731678
LOCUS       CQ731678             2053 bp    DNA             linear      PAT 03-FEB-2004
DEFINITION Sequence 17612 from Patent WO02068579.
ACCESSION   CQ731678
VERSION     CQ731678.1  GI:42308932
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE    1
AUTHORS      Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE        Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
JOURNAL      Patent: WO 02068579-A 17612 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES     source
             1..2053
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      75      Length:      2053
Score:          40.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             6        Gaps:       0

US-10-774-176-10 (1-9) x CQ731678 (1-2053)
Qy 1 PheLeuGlyIleValLeuAlaLeulle 9
Db 1158 TTCTGGGTATTGTTTGTAGCCTGATA 1184

RESULT 16
HS5T4OA
LOCUS       HS5T4OA             2053 bp    RNA             linear      PRI 18-APR-2005
DEFINITION Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.
ACCESSION   Z29083
VERSION     Z29083.1  GI:435654
KEYWORDS    5T4 gene; 5T4 oncofoetal antigen.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE    1 (bases 1 to 2053)
AUTHORS      Myers, K.A., Rahi-Saund, V., Davison, M.D., Young, J.A., Cheater, A.J.
            and Stern, P.L.
TITLE        Isolation of a cDNA encoding 5T4 oncofoetal trophoblast
            glycoprotein. An antigen associated with metastasis contains
            leucine-rich repeats
JOURNAL      J. Biol. Chem. 269 (12), 9319-9324 (1994)
PUBMED      8132670
FEATURES     2 (bases 1 to 2053)
AUTHORS      Myers, K.A.
TITLE        Direct Submission
JOURNAL      Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
            Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
            Location/Qualifiers
             1..2053
               /organism="Homo sapiens"
               /mol_type="other RNA"
               /db_xref="taxon:9606"
               /sex="female"
               /tissue_type="placenta"
               /clone_lib="lambda gt11 library of J. Milan"
               62..372
               /product="LRR N-terminal flank"
               /label="N-Flank"
               85..1347
               /codon_start=1
               /evidence=experimental
               /product="5T4 oncofoetal antigen"
               /protein_id="CAA82324.1"
               /db_xref="GI:435655"
               /db_xref="GOA:Q13641"
               /db_xref="InterPro:IPR000372"
               /db_xref="InterPro:IPR000483"
               /db_xref="InterPro:IPR001611"
               /db_xref="InterPro:IPR003591"
               /db_xref="UniProt/TREMBL:Q13641"
               /translation="MPGCGSRGPAAGDGRLLARLALVLLGWVSSSPTSSASFSSS
               AFLASAVSQPPLPDQCPCSEARATKVCVNRNLTEVPTDLPAYVRLFTGNQ
               LAVLPAGAFARRPLAEALNLGSRDLDEVRAGAFELPSLRDLDSHNPDLADLSPF
               AFGSNASVSPSPPLVELILNHIVPPDERQNRSPFGMVVAALLAGRALQGLRLLELA
               SNHFLYPLDRVLAQLPSLRHLDDLNNLSLVSLTYVSFRNLTHLSLHEDNALKVLHNG
               TLAEQLGPHIRVFLDNNPMWCDLMADMTWLKETEVEVQGRDLTCAYPERQNRVL
               LELNSADLDCDPLPSPSLQTSYVFLGIVLALIGALIFLLVLYLNRGKIKKWHMTDACC
               RDHMEGYHYRYEINADPRLTNLSNSDV"
               130..171
               /sig_peptide
               misc_RNA
               /product="Leucine rich repeat region"
               /label="LRRs"
               966..1119
               /misc_RNA
               /product="LRR C-terminal flank"
               /label="C-flank"
               1153..1215
               /misc_RNA
               /product="transmembrane peptide"
               /standard_name="transmembrane region"
               /function="Anchorage of the protein to the cell membrane"
ORIGIN
Alignment Scores:
Pred. No.:      75      Length:      2053
Score:          40.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0

```

```

Best Local Similarity: 100.0%      Mismatches: 0
Query Match:          100.0%      Indels: 0
DB:                   8           Gaps: 0

US-10-774-176-10 (1-9) x H85T40A (1-2053)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
DB 1156 TTCCTGGGTATTGTTTAGCCCTGATA 1182

RESULT 17
CR855786 2183 bp mRNA linear VRT 03-NOV-2004
LOCUS Xenopus tropicalis finished cDNA, clone TGas020h08.
ACCESSION CR855786
VERSION CR855786.1 GI:55295318
KEYWORDS
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 2183)
AUTHORS Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
Francis,M.D., Garrett,N., Gilchrist,M.J., Grañham,D.V.,
McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
Voigt,J. and Zorn,A.M.
Direct Submission
Submitted (03-NOV-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST/cDNA project.
This sequence is from a Xenopus Gene Collection (XGC) library, from
a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
FEATURES
source
location/Qualifiers
1..2183
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone_lib="XGC-gastrula"
/dev_stage="gastrula (stage 10.5-13 mixed)"

ORIGIN
Alignment Scores:
Pred. No.: 79.5 Length: 2183
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-10 (1-9) x CR855786 (1-2183)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
DB 747 TTTTGGGCATAGTACTAGCCCTCAT 773

RESULT 18
AF063939 2333 bp mRNA linear ROD 01-JAN-2000
LOCUS Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA,
complete cds.
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.

Best Local Similarity: 100.0%      Mismatches: 0
Query Match:          100.0%      Indels: 0
DB:                   8           Gaps: 0

US-10-774-176-10 (1-9) x H85T40A (1-2053)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
DB 1156 TTCCTGGGTATTGTTTAGCCCTGATA 1182

RESULT 17
CR855786 2183 bp mRNA linear VRT 03-NOV-2004
LOCUS Xenopus tropicalis finished cDNA, clone TGas020h08.
ACCESSION CR855786
VERSION CR855786.1 GI:55295318
KEYWORDS
SOURCE Xenopus tropicalis (Silurana tropicalis)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 2183)
AUTHORS Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
Francis,M.D., Garrett,N., Gilchrist,M.J., Grañham,D.V.,
McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
Voigt,J. and Zorn,A.M.
Direct Submission
Submitted (03-NOV-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST/cDNA project.
This sequence is from a Xenopus Gene Collection (XGC) library, from
a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'
end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
FEATURES
source
location/Qualifiers
1..2183
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone_lib="XGC-gastrula"
/dev_stage="gastrula (stage 10.5-13 mixed)"

ORIGIN
Alignment Scores:
Pred. No.: 79.5 Length: 2183
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-10 (1-9) x CR855786 (1-2183)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
DB 747 TTTTGGGCATAGTACTAGCCCTCAT 773

RESULT 18
AF063939 2333 bp mRNA linear ROD 01-JAN-2000
LOCUS Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA,
complete cds.
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2333)
AUTHORS Ninkina,N.N. and Buchman,V.L.
JOURNAL Structure and expression of the rat 5T4 gene
Unpublished
REFERENCE 2 (bases 1 to 2333)
AUTHORS Buchman,V.L.
JOURNAL Direct Submission
Submitted (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK
FEATURES
source
location/Qualifiers
1..2333
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="newborn"
1..2333
/gene="5T4"
1..363
/gene="5T4"
364..1644
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal antigen homolog"
/protein_id="AAF21770.1"
/db_xref="GI:6650212"
/translation="MPGAGSRGSPAGDGRRLRLRLALVILGWVSASAPSSSLPSSSTS
PAFLASGSAQPPPARCPAACESEAAATKCVNRNLLLEVPADLPYPVRLPLTNQ
MTVLPAQAFARQPLADLAVNLGNHLKEVGAGAFELHFGRLDLDSHNPITNLNSAF
TPAGSNVSTPSPLELILNHIVPEQRQNSFGFVAFEGVAAALRSLGLRGL
HHLBLANHFLLPRDLIDLPSLKHLDLRNNSLVSTYASFRNLTHLSLELDNAL
KVLHNSLAEWGGLAHVRVLDNPNVCDVMADVMVSLKETEVPDKARLTCAFPPEK
MNRGLDLTSSDLDCDRTLPSQLTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NIRDARDHMGSHYRIYNADPSLTNLSSSDV"
1645..2333
/gene="5T4"
2315..2320
/gene="5T4"

3'UTR
polyA_signal
ORIGIN
Alignment Scores:
Pred. No.: 84.5 Length: 2333
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-10 (1-9) x AF063939 (1-2333)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
DB 1453 TTCTAGGTATTGTTTAGCTCTGATA 1479

RESULT 19
BD127282 2359 bp DNA linear PAT 18-SRP-2002
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD127282
ACCESSION BD127282
VERSION BD127282.1 GI:23222227
KEYWORDS JP 2002017375-A/2713.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.

```

TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2713 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)
PN JP 2002017375-A/2713
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68/ C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
PC C12P21/02, C12Q1/68/ C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS (424) . . (1572) .

FEATURES
source 1. .2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 85.4 Length: 2359
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-10 (1-9) x BD127282 (1-2359)

Qy 1 PheLeuGlyLeValLeuAlaLeuIle 9
Db 1495 TTCTGGGTATTGTTTAGCCCTGATA 1521

RESULT 20
CQ782724
LOCUS 2359 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2864 from Patent EP1396543.
CQ782724
ACCESSION CQ782724.1 GI:45502667

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.

TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2864 10-MAR-2004;
Research Association for Biotechnology (JP)

FEATURES
source 1. .2359
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"

CDS
LAVPLAPARRPLLAALNLSGKLDVRAFAEHLPLSLQLDLSHPLADLSFP
AFSGSNASVAPSPFLVILLIHVPPEDERQNRSEFGMVVAALLAGRALQGLRLLELA

SNHFLYLPDVLAQLPSLRHLDLNNLSVLSVYFRLTHLESLEHEDNALKVLHNG
TLAELOGLPHIRVFLDNNPWVCHMDMVTWLKETEYVQGKRLTTCAYPEKMRNRVL
LELSADLDCDPILPSPISQTSYVFLGIVLALIGAIFLLVLYLNRKGIKK"

ORIGIN

Alignment Scores: 85.4 Length: 2359
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-10 (1-9) x CQ782724 (1-2359)

Qy 1 PheLeuGlyLeValLeuAlaLeuIle 9
Db 1495 TTCTGGGTATTGTTTAGCCCTGATA 1521

RESULT 21

AK074786
LOCUS 2359 bp mRNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ90305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AK074786
VERSION AK074786.1 GI:22760460
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahari, K., Masubo, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2359)
AUTHORS Isogai, T. and Otsuki, T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp. Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

source

1. .2359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores: 85.4 Length: 2359
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0%

```

Query Match: 100.0%      Indels: 0
DB:          8           Gaps: 0

US-10-774-176-10 (1-9) x AK074786 (1-2359)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
   |||||
Db 1495 TTCTGGGTATTGTTTAGCCCTGATA 1521

RESULT 22
BD127283          2361 bp      DNA      linear      PAT 18-SEP-2002
LOCUS             Primer for synthesizing full-length cDNA and use thereof.
DEFINITION
ACCESSION BD127283
VERSION BD127283.1 GI:23222228
KEYWORDS JP 2002017375-A/2714.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominoidea; Homo.
REFERENCE
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
          Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
          Koga, H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
        PN JP 2002017375-A/2714
        PD 22-JAN-2002
        PP 07-JUL-2000 JP 2000253172
        PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
        PI ISHII,
        PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
        SHINICHI KOJIMA,
        PI TETSUJI OTSUKI, HISASHI KOGA
        PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/00
        PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
        Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers
    1..2361
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 85.5      Length: 2361
Score: 40.00      Matches: 9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match: 100.0%      Indels: 0
DB: 6      Gaps: 0

US-10-774-176-10 (1-9) x CQ782726 (1-2361)
QY 1 PheLeuGlyIleValLeuAlaLeuIle 9
   |||||
Db 1497 TTCTGGGTATTGTTTAGCCCTGATA 1523

RESULT 24
AX961916          2361 bp      DNA      linear      PAT 14-JAN-2004
LOCUS             Sequence 127 from Patent WO03104277.
DEFINITION
ACCESSION AX961916
VERSION AX961916.1 GI:40881326
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominoidea; Homo.
REFERENCE
AUTHORS Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 127 18-DEC-2003;
          Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
    source
    1..2361
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    426..1688
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAF06467.1"
    /db_xref="GI:40881327"
    /translation="MPGCSRGPAAGDGRLELRLALVLLGWSSSPTSSASSFSSS"

CDS
    1..2361
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
    426..1688
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAF06467.1"
    /db_xref="GI:40881327"
    /translation="MPGCSRGPAAGDGRLELRLALVLLGWSSSPTSSASSFSSS"

```

APFLASVAGQPLPDDQPCALCESEAAATVCKVRNLTETVPTDLPAYVRNLFITGNG
LAVLPAGAPAPPEPLAELANLGSRLDEVRAGAFELPSLRQLDLSNPLADLSPP
AFSGNASVSPVVELLNLINIVPPEDRONRSPFGMVVAALLAGRLQGLRLSLA
SNHLYLPDVLPAQLPSLRHLDSNNSLSVLTYSVRNLTLESLEHLENAKLVLENG
TLAEQLGLPHIEVLDPNPMVCDCHMADMTWLKETEVVQGDRLTCAYPEKWRNVL
LEINSAADLDCDILPPLSLQTSYVFLGIVLALGAIFLLVLYLNKRGKKKMMHNRDAC
RDHMEGYHYRYBINADPRUTNLSSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 85.5 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-10 (1-9) x AK961916 (1-2361)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1497 TTCTGGGTATTGTTTAGCCCTGATA 1523

RESULT 25

AK074790

LOCUS Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens 574 oncofetal trophoblast glycoprotein gene.

ACCESSION AK074790.1 GI:22760466

VERSION oligo capping; fis (full insert sequence).

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

Otsuki, T., Ota, T., Nishikawa, T., Hayaashi, K., Suzuki, Y.,
Yamamoto, J., Wakamatsu, A., Kimura, K., Sakamoto, K., Hatano, N.,
Kawai, Y., Ishii, S., Saito, K., Kojima, S., Sugiyama, T., Ono, T.,
Okano, K., Yoshikawa, Y., Aotsuma, S., Sasaki, N., Hattori, A.,
Okumura, K., Nagai, K., Sugano, S. and Isegai, T.
Signal Sequence and Keyword Trap in silico for Selection of
Full-Length Human cDNAs Encoding Secretion or Membrane Proteins
from Oligo-Capped cDNA Libraries
DNA Res. 12, 117-126 (2005)

JOURNAL

REFERENCE

AUTHORS

Isegai, T., Ota, T., Nishikawa, T., Hayaashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahata, K., Masuko, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuma, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.

NEDO human cDNA sequencing project

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
3 (bases 1 to 2361)
Isegai, T. and Otsuki, T.
Direct Submission
Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

source

1. .2361
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="NT2RP2000903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"

/notes="cloning vector: pME18SPL3"

mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 85.5 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x AK074790 (1-2361)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1497 TTCTGGGTATTGTTTAGCCCTGATA 1523

RESULT 26

BC087011

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdepaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 186 Row: 0 Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

source

1. .2361
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clones="MGC:93332 IMAGE:7193411"
 /tissue types="Heart, rat (Brown Norway)"
 /clone_lib="NIH MGC_234"
 /lab_host="DH10B"
 /note="vector: pExpress1"

gene

1. .2361
 /gene="Tpbp"
 /note="synonym: 5T4"
 /db_xref="GeneID:83684"
 /db_xref="RGD:621453"
 364..1644

CDS

/gene="Tpbp"
 /codon_start=1
 /product="Tpbp protein"
 /protein_id="AAH87011.1"
 /db_xref="GI:56268820"
 /db_xref="GeneID:83684"
 /db_xref="RGD:621453"
 /translation="MPCAGSRGSPAGDGLRLAELALVLLGMVSASAPSSLSSTNS
 PAAPLAGSQAPPPACRACCESEARTVKVNRNLLLEVPADLPYVNRFLFTGNQ
 MVLVPAFAGAPQLADLAVLGNHLKEVGAFHPLGRLDLSHNPLTFLSAP
 TPAAGSNVSTPQPLILNLHIVPPEDQONGSFEQGVAFEGMVAALRSGALRGL
 HLELASNHYLYLPRLDLQPLSLKHLDRNLSLVSTYASFRNLTHLESLSHLDNAL
 KVLHNSFLASQGLAHVFLDNNPWCDCVMADMVSLKETEVPVDPKARLTCAPEK
 MNRGLDLTSSDLDCDATALPQSLQTSYVFLGIVLALIGLFLVLYLNKRGIKKWH
 NIRDACHDMEGHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
 Pred. No.: 85.5 Length: 2361
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-10 (1-9) x BC087011 (1-2361)

QY 1 PheLeuGlyIleValLeuAlaLeuIle 9

|||||
 1453 TTCTAGGATATGTTTAGCTCTGATA 1479

RESULT 27

BC037161

LOCUS

DEFINITION BC037161 2379 bp mRNA linear PRI 29-JUN-2004
 IMAGE:4138906), complete cds.

ACCESSION

BC037161

VERSION

BC037161.2

KEYWORDS

MGC.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 2379)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Carninci, P., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smal, D.B.,
 Sutterich, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2379)

Strausberg, R.

Direct Submission

Submitted (03-SEP-2002)

National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 19, 2003 this sequence version replaced gi:22713382.

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>Contact: nisc.mc@nhgri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Iaric, J., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAX Plate: 26 Row: m Column: 15

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 5729717.

FEATURES

source

1. .2379
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="MGC:15317 IMAGE:4138906"
 /tissue type="Muscle, rhabdomyosarcoma"
 /clone_lib="NIH MGC_17"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"
 1. .2379
 /gene="TPBG"
 /note="synonym: M6P1, 5T4-AG, 5T4"
 /db_xref="GeneID:7162"
 /db_xref="MIM:190920"
 427..1689
 /gene="TPBG"
 /codon_start=1
 /product="5T4 oncofetal trophoblast glycoprotein"

gene

CDS

```

/protein_id="AAH37161.1"
/db_xref="GI:22713383"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"
/translation="MPGSGRGAAGDGRRLRLARLALVLLGWSSSPTSSASSFSSS
APFLASAVAQPPPLPQCPALCSBAARTVKCNRLNTEVPTDLPAYVRLNLTGNO
LAVLPAGAPARQPPPLADLEALNLSGNHLKEVCAGAFELPLGLRLDLSHNLTLNLSA
AFAGSNASVASPGLLEELIINHIVPPDQKQNGFEGWAFEGWVAALRGLALRGL
SNHFLPSNHFPLPRDLAQPSRLYDLRNNLSVLTYSFRLTHLESJHLEDNAL
TLAELQGLPHIRVFLDNNPWCDCMADMTVTLKETEVQGDRLTCAYPEKVRNVL
LELNSADLCDDPILPSPQLTSYVFLGIVLALIGAIFLLVLYLNKGIKKWMMHNRD
RDHMEGYHYRYEINADPRLNLSNSDV"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 86.1 Length: 2379
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

```

US-10-774-176-10 (1-9) x BC037161 (1-2379)

```

QY 1 PheLeuGlyIleValLeuAlaLeulle 9
|||||
Db 1498 TTCCTGGGTATTTTAGCCCTGAYA 1524

```

RESULT 28

```

BC058198
LOCUS BC058198 2423 bp mRNA linear ROD 21-OCT-2003
DEFINITION Mus musculus trophoblast glycoprotein, mRNA (cDNA clone MGC:68145
IMAGE:5353871), complete cds.

```

```

ACCESSION BC058198
VERSION BC058198.1 GI:34849573
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)

```

ORGANISM

```

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Rutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

```

REFERENCE

```

1 (bases 1 to 2423)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaeetz, T.E., Brownstein, M.J., Usdin, T.B., Toehiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywiński, M.I., Skaleks, U., Snailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2423)
Strausberg, R.

```

Direct Submission

```

Submitted (15-SEP-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk

```

REMARK

COMMENT

Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hghri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, F., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maekeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tougeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 123 Row: p Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

source

```

1. .2423
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:68145 IMAGE:5353871"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal Carcinoma, 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"

```

gene

```

1. .2423
/gene="Tpbq"
/notes="synonym: 5T4"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
402. .1682
/gene="Tpbq"
/codon_start=1
/product="Tpbq protein"
/protein_id="AAH58198.1"
/db_xref="GI:34849574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
/translation="MPGAGSGRGAAGDGRRLRLARLALVLLGWSSASAPSSSPSSSTS
PAAPFLASGSAQPPPAERCPAACSCSEARTVKCNRLNTEVPTDLPAYVRLNLTGNO
MTVLPAAGAPARQPPPLADLEALNLSGNHLKEVCAGAFELPLGLRLDLSHNLTLNLSA
AFAGSNASVASPGLLEELIINHIVPPDQKQNGFEGWAFEGWVAALRGLALRGL
KVLHNSFLAEGWGLAHVKVFLDNNPWCDCMADMTVTLKETEVQGDRLTCAYPEKARLTCA
MRNGGLDLSNSDLDCCDAVLQPSLQTSYVFLGIVLALIGAIFLLVLYLNKGIKKWMMH
NIRDACDRDHMEGYHYRYEINADPRLNLSNSDV"
642. .1262

```

CDS

```

1. .2423
/gene="Tpbq"
/notes="synonym: 5T4"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
402. .1682
/gene="Tpbq"
/codon_start=1
/product="Tpbq protein"
/protein_id="AAH58198.1"
/db_xref="GI:34849574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
/translation="MPGAGSGRGAAGDGRRLRLARLALVLLGWSSASAPSSSPSSSTS
PAAPFLASGSAQPPPAERCPAACSCSEARTVKCNRLNTEVPTDLPAYVRLNLTGNO
MTVLPAAGAPARQPPPLADLEALNLSGNHLKEVCAGAFELPLGLRLDLSHNLTLNLSA
AFAGSNASVASPGLLEELIINHIVPPDQKQNGFEGWAFEGWVAALRGLALRGL
KVLHNSFLAEGWGLAHVKVFLDNNPWCDCMADMTVTLKETEVQGDRLTCAYPEKARLTCA
MRNGGLDLSNSDLDCCDAVLQPSLQTSYVFLGIVLALIGAIFLLVLYLNKGIKKWMMH
NIRDACDRDHMEGYHYRYEINADPRLNLSNSDV"
642. .1262

```

misc_feature

```

/gene="Tpbq"
/notes="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [Function unknown]"
/db_xref="CDD:COG4886"
1299. .1415
/gene="Tpbq"
/notes="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

```

misc_feature

```

/gene="Tpbq"
/notes="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 87.6 Length: 2423
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

```



```

Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-10 (1-9) x BC058198 (1-2423)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1491 TTCTAGGTATTGTTTGTAGCTCTGATA 1517

RESULT 29
AX961912 2557 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 123 from Patent WO03104277.
ACCESSION AX961912
VERSION AX961912.1 GI:40881322
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE
1 Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
Stat6 activation gene
Patent: WO 03104277-A 123 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)

FEATURES
source
Location/Qualifiers
1. .2557
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAFO6465.1"
/db_xref="GI:40881322"
/translation="MPGAGSGPSAGDGRRLRLARLALVLGWVSASAPSSVSSSTS
PADFLASGSAQPPPAECSEAAATVKVNRNLLVPPADLPYPYVRNPLFTGNQ
MTVLPAGAFARQPLADLEALNLSGHLKEVCAGAFHLPGLRLDLSHNPLTNLSAF
VPAGSNASVAPSLBELILNHI VPPEDQONGSFGWVAFEGMVAALRSLGLALRGL
TRLEASNHFLPLPRDLIAQLPSLYLDLNNLSVSLTYASFNLTLSLHLEDNAL
KVLNLSLAEWQGLAHVKVFLDNNPWCDCMADWAMVKETEVPVDPKARLTCAPPEK
MNRGLDLNSSLDCDAVLPSQTSYVFLGIVLALIGAIFLLVLYLNRRKGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNSDV"

ORIGIN
Alignment Scores: 92.1 Length: 2557
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

CDS
US-10-774-176-10 (1-9) x AX961914 (1-2557)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 1645 TTCTAGGTATTGTTTGTAGCTCTGATA 1671

RESULT 31
AX961914 2714 bp mRNA linear PRI 18-JUN-2005
LOCUS
DEFINITION Macaca fascicularis testis cDNA clone: Qcsa-11109, similar to human
trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.
ACCESSION AX961914
VERSION AX961914.1 GI:67967899
KEYWORDS
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE
1 International consortium for macaque cDNA sequencing and analysis.
DNA sequences of macaque genes expressed in brain or testis and its
evolutionary implications
Unpublished

JOURNAL
REFERENCE
2 Osada, N., Hirata, M., Tanuma, R., Kusuda, J., Hida, M., Suzuki, Y.,
Sugano, S., Gojobori, T., Shen, J.-C.-K., Wu, C.-I. and Hashimoto, K.
Substitution rate and structural divergence of 5'UTR evolution:
Comparative analysis between human and cynomolgus monkey cDNAs
Unpublished

JOURNAL
REFERENCE
3 (bases 1 to 2714)
Hashimoto, K., Kusuda, J. and Sugano, S.
Direct Submission
Submitted (18-MAR-2004) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
The International consortium for macaque cDNA sequencing and
analysis consists of: Department of Virology and Human Genome
Center, Institute of Medical Science, The University of Tokyo.

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```



```

Alignment Scores:
Pred. No.: 190 Length: 5551
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x HSA012159 (1-5551)

Qy 1 PheLeuGlyIleValLeuAlaLeuile 9
|||||
Db 4502 TTCTGGGTATTGTTTGTAGCCCTGATA 4528

RESULT 33
MMU012160
LOCUS
DEFINITION Mus musculus 574 oncofetal trophoblast glycoprotein gene.
ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 574 gene; 574 oncofetal trophoblast glycoprotein.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
King, K.W., Sheppard, P.C., Westwater, C., Stern, P.L. and Myers, K.A.
Organisation of the mouse and human 574 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
Biochim. Biophys. Acta 1445 (3), 257-270 (1999)

JOURNAL
PUBMED 10366710
REFERENCE 2 (bases 1 to 7942)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK

FEATURES
source
1..7942
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="Lambda Dash"
3108..3113
/bound_moiety="Sp1"
misc_binding
3114..3119
/bound_moiety="Sp1"
misc_binding
3124..5779
/gene="574"
3124..3151
/gene="574"
3152..3450
/gene="574"
3451..5779
/gene="574"
3779..5059
/gene="574"
/codon_start=1
/product="574 oncofetal trophoblast glycoprotein"
/protein_id="CAA09931.1"
/db_xref="GI:3805949"
/db_xref="GOA:Q9Z0L0"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProt/TREMBL:Q9Z0L0"
/translation="MPGAGRGPSAGDGRRLRLARLALLVLLGWVSASAPSSVPSSTYS
PADPLASGSQPPPAERCPAACESEAARTVKVNRNLLLEVPADLPYPVRLFLTGNG

MTVLPAGAPARQPPLADLEALNLSGNHLKEVCAGAFEHLPGRLRLDLSHNPITNLSAF
VPAGNSASVSPLEELILNHI VPPEDORONGSPFGWVAFPGWVAALRSGLALRGL
TRLELASNHPLFLPRDLIAQLPSLAYLDRNNLSLYTASFRNLTHLESLHEDNAL
KVLNRSTLAEWGGLAHVFLDNNPWCCTPADWVAMLKETEVPDPKARLCAFFPK
MNRGLDLNSSLDCDCAVLPSQLQTSYVFLGIVLALIGAIPLLVLVLRKGIKKWMH
NIRDACRDHMEGYHYREINADPRLTNLSSSDV"
sig_peptide
3779..3865
/mat_peptide
3866..5056
/gene="574"
3866..5056
/gene="574"
/product="574 oncofetal trophoblast glycoprotein"
5713..5718
/polyA_signal
5759..5764
/polyA_signal
5759..5764
/gene="574"
5759..5764
/gene="574"

ORIGIN
Alignment Scores:
Pred. No.: 266 Length: 7942
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-10 (1-9) x MMU012160 (1-7942)

Qy 1 PheLeuGlyIleValLeuAlaLeuile 9
|||||
Db 4868 TTCTAGGTATTGTTTGTAGCTGATA 4894

RESULT 34
HSJ492P14
LOCUS
DEFINITION Human DNA sequence from clone RP3-492P14 on chromosome 6q13-15
Contains a single stranded DNA binding protein pseudogene, the TPBG
gene for trophoblast glycoprotein (574-AG) and a CpG island,
complete sequence.
ACCESSION AL121977
VERSION AL121977.11 GI:11863678
KEYWORDS HTG; CpG island; TPBG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 121909)
REFERENCE 1
AUTHORS Garner, P.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Dec 15, 2000 this sequence version replaced gi:11558491.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP3-492P14 is from the library RPCI-3 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all

```

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

```

source
  1. .121909
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="6"
    /map="q13-15"
    /clone_lib="RP3-492P14"
    /clone="RP3-492P14"
    /clone="RP3-492P14"
    /clone="RP3-492P14"
    /note="Clone right end: RP1-93K22"
    complement(10004..10982)
    /locus_tag="RP3-492P14.2-001"
    /pseudo
    complement(10004..10982)
    /locus_tag="RP3-492P14.2-001"
    /note="match: proteins: P81877 Q99LX9 Q9BWW6 Q9CYZ8 Q9D6L4 Q9P038 Q9I477"
    /pseudo
    /codon_start=1
    86539
    /note="Clone left end: RP1-90G1"
    /genes="TPBG"
    109639..116836
    /locus_tag="RP3-492P14.1-001"
    join(109639..109916,110631..116836)
    /genes="TPBG"
    /locus_tag="RP3-492P14.1-001"
    /product="trophoblast glycoprotein"
    /note="match: ESTs: AA149121 AA152323 AA565852 AA643734 AL544610 AW471072 AW662538 BE260089 BF306457 BF306926 BF314984 BI196133 BI562387 BM069633 BM670613 match: cDNAs: AJ420536.1 Z29083.1"
    110970..112232
    /genes="TPBG"
    /locus_tag="RP3-492P14.1-001"
    /standard_name="OTTHUMP0000016786"
    /note="match: proteins: Q13641 Q9QZD9 Q9ZOL0"
    /codon_start=1
    /product="trophoblast glycoprotein"
    /protein_id="CA121546.1"
    /db_xref="GI:56203539"
    /db_xref="GenBank:12004"
    /db_xref="GOA:Q13641"
    /db_xref="InterPro:IPR000372"
    /db_xref="InterPro:IPR000483"
    /db_xref="InterPro:IPR001611"
    /db_xref="InterPro:IPR003591"
    /db_xref="UniProt/TREMBL:Q13641"
    /translation="MPGSGRGPAAGDGRLEALRLALVLLGWSSSPTSSASSPSS
    APLASAVAQPLPQPCALCESEARTYKCNRLTEVPTDLPAYVNLRLTGNG
    LAVLPAGARFPPLVFLNLNHTVPPEDQRNRQFVAGLHPLSLQDLNHLADLSFF
    AFGSNAGVAFSPVVELNLNHTVPPEDQRNRQFVAGLHPLSLQDLNHLADLSFF
    SHNLYLPDLQLSLDLNLSNLSVLTYSFRLNTHLSLEHEDNALKVLHNG
    TLAEGLGPHIRVFLDNNPWCDCMADMTWLKTEVYVQGDRLTCAVPEKRNRLV
    BLN8ADLDCDPTLP8LQTSVYVLGIVLALIGALIFLLVLYLNKGIKKWMMNIRDAC
    RDHMGYHYRIENADPLRNLSNSDV"
    116817..116822
    /genes="TPBG"
    /locus_tag="RP3-492P14.1-001"
    116836
    /genes="TPBG"
    /locus_tag="RP3-492P14.1-001"
    /note="Clone right end: RP3-492P14"
  
```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.41e+03 Length: 121909
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-10 (1-9) x HSJ492P14 (1-121909)
Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 112041 TTCTGGGTATTGTTTAGCCCTGATA 112067

RESULT 35
LOCUS BX248521 144888 bp DNA linear VRT 29-APR-2004
DEFINITION Zebrafish DNA sequence from clone CH211-243C14 in linkage group 23, complete sequence.
ACCESSION BX248521
VERSION BX248521.12 GI:46878960
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
            Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
            1 (bases 1 to 144888)
            Clark,S.
            Direct Submission
            Submitted (29-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Apr 29, 2004 this sequence version replaced gi:46406548.
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -----
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WormPEP; Information
            on the WormPEP database can be found at
            http://www.sanger.ac.uk/Projects/c_elegans/wormpep/Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU). For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-243C14 is from a CHORI-211 BAC library
            VECTOR: PTARBAC2.1.
            Location/Qualifiers
            1..144888
            /organism="Danio rerio"
  
```

FEATURES

source

/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-243C14"
/clone_lib="CHORI-211"

ORIGIN

Alignment Scores:
Pred. No.: 4e+03 Length: 144888
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-10 (1-9) x BX248521 (1-144888)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 50025 TTCCTTGGTATTGTTTGGCTTGGATT 50051

RESULT 36

AC158516/c
LOCUS AC158516 167046 bp DNA linear ROD 21-JUN-2005
DEFINITION Mus musculus BAC clone RP24-511A23 from chromosome 9, complete sequence.

ACCESSION AC158516 AC117768
VERSION AC158516.2 GI:63025421

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 167046)
Adams, S., Cotton, M. and Haglund, K.

AUTHORS

The sequence of Mus musculus BAC clone RP24-511A23

JOURNAL

Unpublished (2001)

REFERENCE

2 (bases 1 to 167046)
Wilson, R.K.

AUTHORS

Direct Submission

JOURNAL

Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

REFERENCE

3 (bases 1 to 167046)
Wilson, R.K.

AUTHORS

Direct Submission

JOURNAL

Submitted (21-JUN-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

AUTHORS

On May 4, 2005 this sequence version replaced gi:61656412.

COMMENT

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@genome.wustl.edu

----- Summary Statistics

Center project name: M_BR0511A23

Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e. phred quality
>=30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone, fosmid clone or direct clone walk sequence.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to obtain the consensus sequence. The

assembly was confirmed by restriction digest.

This finishing standard has slightly changed from the previous
Human standard. Specifically, standards for regions of low sequence
complexity (such as dinucleotide repeats and small unit tandem
repeats) have been relaxed. These regions are very prevalent in the
mouse genome, and the return on extended finishing efforts is
minimal.

If a sequence meets the criteria of the above statement, it needs
no comments or tags. If the criteria are not met, such as ambiguous
bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:

The BAC Library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

This sequence is the entire insert of the clone.

FEATURES

source

1..167046
/location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/clone="RP24-511A23"
/clone_lib="RPCI-24"
16685..16712
/note="Sequence derived from PCR product of genomic DNA"

misc_feature

31565..31779
/note="Unresolved simple sequence repeat."

unseq

46721..46808
/note="Unresolved simple sequence repeat."

unseq

142336..142347
/note="Sequence derived from one plasmid subclone."

ORIGIN

Alignment Scores:
Pred. No.: 4.57e+03 Length: 167046
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-10 (1-9) x AC158516 (1-167046)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
Db 109749 TTCCTAGGTATTGTTAGCTCTGATA 109723

RESULT 37

LOCUS

AC128294/c

DEFINITION

Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.

ACCESSION

AC128294

VERSION

AC128294.3 GI:25083347

KEYWORDS

HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
1 (bases 1 to 210237)
Muzny, D., Warle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Crase, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falle, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulvik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshuwa, L., Loulesed, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoc, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, C., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Soderstrom, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

TITLE JOURNAL REFERENCE AUTHORS

2 (bases 1 to 210237)

TITLE JOURNAL

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 210237)

Rat Genome Sequencing Consortium.

TITLE JOURNAL

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 19, 2002 this sequence version replaced gi:23265004. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GZGV
Center clone name: CH230-176H20
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 201781 bases at least Q40
Consensus quality: 203921 bases at least Q30
Consensus quality: 205310 bases at least Q20
Estimated insert size: 205531; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 210237: contig of 210237 bp in length.

FEATURES

source

1. 210237
/organism="Rattus norvegicus"
/db_type="genomic DNA"
/mol_xref="taxon:10116"
/clone="CH230-176H20"

misc_feature

1. 1142
/note="wgs_end_extension
clone_end:T7"

misc_feature

2177. 144799
/note="clone_boundary
clone_end:T7
site:

misc_feature

end sequence:BH360464"
complement(206062..206961)
/note="clone_boundary
clone_end:Sp6
site:

misc_feature

end sequence:BH360465"
208907..210237
/note="wgs_end_extension
clone_end:Sp6"

ORIGIN

Alignment Scores:

Pred. No.:	5.67e+03	Length:	210237
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	14	Gaps:	0

US-10-774-176-10 (1-9) x AC128294 (1-210237)

Qy

1 PhaeuclyleValleuAlaLeuile 9
|||||

Db 110429 TTCTAGGTATTGTTTAGCTCGATA 110403

RESULT 38

AC106962/c

LOCUS

DEFINITION

AC106962

VERSION

KEYWORDS

AC106962 239076 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC106962
AC106962.5 GI:25139469
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 239076)
AUTHORS	Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angulano,D., Ayalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,B., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleeland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.G., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guvera,W., Gunnaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hayes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,L., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,B., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kellay,S., Khan,Z., King,L., Kovac,C., Kowicz,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensheua,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokeme,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,P., Polinder,A., Popovic,D., Primus,E., Pu,L.-L., Puaro,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A. Direct Submission
TITLE	Unpublished
JOURNAL	2 (bases 1 to 239076)
REFERENCE	Worley,K.C.
AUTHORS	Direct Submission
TITLE	Submitted (14-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	3 (bases 1 to 239076)
REFERENCE	Rat Genome Sequencing Consortium.
AUTHORS	Direct Submission
TITLE	Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	On Nov 20, 2002 this sequence version replaced gi:22857070. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
COMMENT	

DEFINITION Bos taurus clone CH240-172A22, *** SEQUENCING IN PROGRESS ***, 21
unordered pieces.
ACCESSION AC165904
VERSION AC165904.1 GI:70980581
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS
Muzny, D., Adams, C., Agbai II, O., Allen, C., Albrooks, S., Archer, P., Arrédon, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R., Berarducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P., Cadoree, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, J., Cierck, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, M., Holder, M., Haaland, W., Haebler, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B., Howell, L. T., Hui, S., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S., Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B., Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorenshuwa, L., Lozano, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, S., McClelland, H., McPherson, J., Mercadado, C., Metzker, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastough, E., Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puzo, M., Qin, X., Quinn, A., Quiroz, J., Rabata, D., Rachlin, E., Raigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, P., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savary, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R., Trejos, Z., Usmani, K., Varago, C., Verduzco, D., Villasana, D., Virk, D., Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, P., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 66940)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help.tmc.edu
----- Project Information
Center project name: FKIS
Center clone name: CH240-172A22
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 93363 bases at least Q40

Consensus quality: 97762 bases at least Q30
Consensus quality: 100227 bases at least Q20
Estimated insert size: 106512; sum-of-contigs estimation
Quality coverage: 1x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
3372: contig of 3372 bp in length
3472: gap of unknown length
6293: contig of 2821 bp in length
6393: gap of unknown length
9161: contig of 2768 bp in length
9261: gap of unknown length
11790: contig of 2529 bp in length
11890: gap of unknown length
16222: contig of 4332 bp in length
16322: gap of unknown length
19241: contig of 2919 bp in length
19341: gap of unknown length
21524: contig of 2183 bp in length
21624: gap of unknown length
25347: gap of unknown length
2806: contig of 3359 bp in length
2806: gap of unknown length
32333: contig of 3527 bp in length
32433: gap of unknown length
35821: contig of 3388 bp in length
35921: gap of unknown length
40357: contig of 4436 bp in length
40457: gap of unknown length
45456: contig of 4999 bp in length
45556: gap of unknown length
47778: contig of 2222 bp in length
47778: gap of unknown length
50273: contig of 2395 bp in length
50373: gap of unknown length
52465: contig of 2092 bp in length
52465: gap of unknown length
55477: contig of 2912 bp in length
55577: gap of unknown length
59739: contig of 4162 bp in length
59839: gap of unknown length
62176: contig of 2337 bp in length
62276: gap of unknown length
64571: contig of 2295 bp in length
64572: gap of unknown length
64672: contig of 2269 bp in length.
Location/Qualifiers
1. .66940
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-172A22"
3372..3472
/estimated_length=unknown
6294..6393
/estimated_length=unknown
9162..9261
/estimated_length=unknown
11791..11890
/estimated_length=unknown
16223..16322
/estimated_length=unknown
19242..19341

FEATURES
source
gap
gap
gap
gap
gap
gap

gap /estimated_length=unknown
 21525..21624
 /estimated_length=unknown
 25248..25347
 /estimated_length=unknown
 28707..28806
 /estimated_length=unknown
 32334..32433
 /estimated_length=unknown
 35822..35921
 /estimated_length=unknown
 40358..40457
 /estimated_length=unknown
 45457..45556
 /estimated_length=unknown
 47779..47878
 /estimated_length=unknown
 50274..50373
 /estimated_length=unknown
 52466..52565
 /estimated_length=unknown
 55478..55577
 /estimated_length=unknown
 59740..59839
 /estimated_length=unknown
 62177..62276
 /estimated_length=unknown
 64572..64671
 /estimated_length=unknown

ORIGIN

Alignment Scores:

Pred. No.: 3.29e+03 Length: 66940
 Score: 39.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 97.5% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-10 (1-9) x AC165904 (1-66940)

Qy 1 PheLeuGlylleValleuAlaLeuille 9

Db 36784 TTCTTAGGGTGTCTTTGGGGTTAATC 36810

RESULT 40

AC150616 174454 bp DNA linear HTG 28-JUL-2004
 LOCUS Callithrix jacchus clone CH259-471J6, WORKING DRAFT SEQUENCE, 3
 DEFINITION unordered pieces.

ACCESSION AC150616

VERSION AC150616.1 GI:50726770

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Callithrix jacchus (white-tufted-ear marmoset)

ORGANISM

Callithrix jacchus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
 Callitrichidae; Callithrix.

REFERENCE

AUTHORS 1 (bases 1 to 174454)
 Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
 Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
 Coleman,H., Daki,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
 Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
 Karlins,E., Kim,H., Kwong,P., Laric,P., Larson,S., Lee-lin,S.-Q.,
 Legaspi,R., Madden,M., Maduro,Q.B., Maduro,V.B., Margulies,E.H.,
 Masiello,C., Maskeri,B., McDowell,J., Mullikin,J.C., Paquinigan,C.,
 Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddix-Dugue,N.,
 Schandler,K., Schueler,M.G., Shah,K., Sison,C., Stantiripop,S.,
 Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D.,
 Young,A. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 174454)

TITLE

NISC Comparative Sequencing Initiative

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 174454)

AUTHORS

Green,E.D.
 Direct Submission
 Submitted (28-JUL-2004) NIH Intramural Sequencing Center, 8717
 JOURNAL GroveMont Circle, Gaithersburg, MD 20877, USA

COMMENT

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: hlb

Center clone name: 471J06

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 173676 bases at least Q40

Consensus quality: 173887 bases at least Q30

Consensus quality: 173977 bases at least Q20

Insert size: 178000; agarose-fp

Insert size: 174254; sum-of-contigs

Quality coverage: 7.84x in Q20 bases; agarose-fp

Quality coverage: 8.01x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 43236: contig of 43236 bp in length

* 43237 43236: gap of unknown length

* 43337 99626: contig of 56290 bp in length

* 99627 99726: gap of unknown length

* 99727 174454: contig of 74728 bp in length.

FEATURES

source

1..174454
 /location/Qualifiers
 /organism="Callithrix jacchus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9483"
 /clone="CH259-471J6"
 /clone_lib="CH259"
 /note="BAC resource: http://bacpac.chori.org/"

misc_feature

1..43236
 /note="assembly_fragment"

clone_end:17

vector_side:left

43237..43336

/estimated_length=unknown

gap

43337..99626

/note="assembly_fragment"

clone_end:SP6

vector_side:right

99627..99726

/estimated_length=unknown

99727..174454

/note="assembly_fragment"

ORIGIN

Alignment Scores:

Pred. No.: 8.06e+03 Length: 174454
 Score: 39.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 97.5% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-10 (1-9) x AC150616 (1-174454)

Qy

1 PheLeuGlylleValleuAlaLeuille 9

|||||||:|||||||

```

Db      6457 TTTTGGGTGTTGTTCTTCTTAAAT 6483

RESULT 41
HSPA32B9/c
LOCUS   HSPA32B9          421 bp    DNA        linear    STS 21-MAY-1998
DEFINITION H.sapiens flow-sorted chromosome 6 HindIII fragment, SC6pA32B9,
sequence tagged site.
ACCESSION Z94208
VERSION   294208.1 GI:1945202
KEYWORDS STS; single read.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 421)
Mungall,A.J., Huckle,E., Langford,C., Ross,M.T. and Hunt,S.E.
Direct Submission
TITLE     Submitted (17-APR-1997) The Sanger Centre, Wellcome Trust Genome
JOURNAL   Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
humquery@sanger.ac.uk
COMMENT   Vector: pBSISK+.
FEATURES             Location/Qualifiers
     source           1..421
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="SC6pA32B9"
                     /sex="female"
                     /tissue type="EBV lymphoblastoid cell line"
                     /clone_lib="SC6pA"
                     /dev_stage="adult"
                     /notes="The estimated purity of the flow-sorted chromosome
                     6 library is >97%"

ORIGIN
Alignment Scores:
Pred. No.:      48.9      Length:      421
Score:          38.00     Matches:      8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match:    95.0%   Indels:      0
DB:             10      Gaps:       0

US-10-774-176-10 (1-9) x HSPA32B9 (1-421)

Qy      1 PheLeuGlyIleValLeuAlaLeuIle 9
Db      241 TTCTGGGTATGTTTGTAGCCATGATA 215

RESULT 42
AU026869
LOCUS   Rattus norvegicus, OTSUKA clone, OT23.18/093B02, microsatellite
DEFINITION Rattus norvegicus, sequence tagged site.
ACCESSION AU026869
VERSION   AU026869.1 GI:4516792
KEYWORDS STS.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (sites)
Watanabe,T.K., Hishigaki,H., Kanemoto,N., Miyakita-Mizoguchi,A.,
Oga,K., Okuno,S., Ono,T., Tsuji,A., Hayashi,H., Adachi,M.,
Yamasaki,Y., Iriye,Y., Takahashi,E., Takagi,T., Nakamura,Y. and
Tanigami,A.
The large-scale mapping of rat microsatellite markers
Unpublished
REFERENCE 2 (bases 1 to 440)
Watanabe,T.K.

TITLE     Direct Submission
JOURNAL   Submitted (24-JUL-1998) Takeshi K Watanabe, Otsuka Pharmaceutical
Co., Ltd., Otsuka GEN Research Institute; 463-10, Kagasuno,
Kawauchi-cho, Tokushima, Tokushima 771-0192, Japan
(E-mail:watanabe@otsuka.genome.ad.jp, Tel:81-886-65-2888,
Fax:81-886-37-1035)
FEATURES             Location/Qualifiers
     source           1..440
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="OT23.18/093B02"
                     /notes="OT23.18/093B02F-5'-AGCAGTTTCTAGAACCCGT-3',
                     OT23.18/093B02R-5'-TATGACACACACATACCTGCT-3'"

ORIGIN
Alignment Scores:
Pred. No.:      51      Length:      440
Score:          38.00     Matches:      8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match:    95.0%   Indels:      0
DB:             10      Gaps:       0

US-10-774-176-10 (1-9) x AU026869 (1-440)

Qy      1 PheLeuGlyIleValLeuAlaLeuIle 9
Db      12 TTCTTGGGTATGTTTGTAGCTCTCTC 38

RESULT 43
AL807766/c
LOCUS   AL807766          58433 bp    DNA        linear    PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP3-383J8 on chromosome X, complete
sequence.
ACCESSION AL807766
VERSION   AL807766.4 GI:21738718
KEYWORDS HTG.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 58433)
Heath,P.
Direct Submission
TITLE     Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jul 12, 2002 this sequence version replaced gi:21727592.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP3-383J8 is from the library RPCI-3 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such

```

as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

```

source
1. .58433
  /location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /clone="RP3-38338"
  /clone_lib="RPCI-3"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 4.91e+03 Length: 58433
Score: 38.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.0% Indels: 0
DB: 8 Gaps: 0

```

US-10-774-176-10 (1-9) x AL807766 (1-58433)

```

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 45420 TTCCTAGGATAGTACTAGCCCTTC 45394

```

RESULT 44

AL954182/c

LOCUS AL954182 151552 bp DNA linear VRT 25-NOV-2003
 DEFINITION Zebrafish DNA sequence from clone CH211-205N18 in linkage group 20, complete sequence.

ACCESSION AL954182

VERSION AL954182.18 GI:38521360

KEYWORDS HTG.

SOURCE Danio rerio (zebrafish)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 151552)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
 Submitted (25-NOV-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk
 On Nov 25, 2003 this sequence version replaced gi:38230013.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where

this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
 CH211-205N18 is from a CHORI-211 BAC library
 VECTOR: pIARBAC2.1.

FEATURES

source

```

1. .151552
  /location/Qualifiers
  /organism="Danio rerio"
  /mol_type="genomic DNA"
  /db_xref="taxon:7955"
  /clone="CH211-205N18"
  /clone_lib="CHORI-211"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.2e+04 Length: 151552
Score: 38.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.0% Indels: 0
DB: 5 Gaps: 0

```

US-10-774-176-10 (1-9) x AL954182 (1-151552)

```

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9
|||||
Db 33595 TTCCTAGGATAGTACTAGCCCTTC 33569

```

RESULT 45

BX005450/c

LOCUS BX005450 164218 bp DNA linear VRT 30-NOV-2003

DEFINITION Zebrafish DNA sequence from clone CH211-150Q23, complete sequence.

ACCESSION BX005450

VERSION BX005450.9 GI:38174975

KEYWORDS HTG.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 164218)

Smith, M.

Direct Submission

Submitted (29-NOV-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk

On Nov 4, 2003 this sequence version replaced gi:37665469.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where

Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-150023 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES

source
Location/Qualifiers
1..164218
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-150023"
/clone_lib="CHORI-211"

ORIGIN

Alignment Scores:
Pred. No.: 1.29e+04 Length: 164218
Score: 38.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-10 (1-9) x BX005450 (1-164218)

Qy 1 PheLeuGlyLeValLeuAlaLeuIle 9

Db 50134 TTCTTGGTATTGTTTGTAGCTTTCCTT 50108

RESULT 46

LOCUS BX323031 164748 bp DNA linear VRT 24-AUG-2004
DEFINITION Zebrafish DNA sequence from clone DKEYP-7B3 in linkage group 18, complete sequence.

ACCESSION BX323031

VERSION BX323031.7 GI:51035888

KEYWORDS HTG.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 164748)

Sycamore, N.

Direct Submission

Submitted (24-AUG-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 6, 2004 this sequence version replaced gi:50949643.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEYP-7B3 is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5.

FEATURES

source
Location/Qualifiers
1..164748
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEYP-7B3"
/clone_lib="DanioKeyPilot"

ORIGIN

Alignment Scores:
Pred. No.: 1.29e+04 Length: 164748
Score: 38.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-10 (1-9) x BX323031 (1-164748)

Qy 1 PheLeuGlyLeValLeuAlaLeuIle 9

Db 94412 TTCTCGGTATTGTTTGTAGCTTTCCTT 94438

RESULT 47

LOCUS CR855265 166117 bp DNA linear HTG 05-NOV-2004
DEFINITION Danio rerio clone CH211-195E6, *** SEQUENCING IN PROGRESS ***, 25 unordered pieces.

ACCESSION CR855265

VERSION CR855265.3 GI:55468451

KEYWORDS HTG; HTGS PHASE1.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 166117)

McLay, K.

Direct Submission

Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 5, 2004 this sequence version replaced gi:55250949.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Web site: <http://www.sanger.ac.uk>

Contact: zfish-help@sanger.ac.uk

----- Project Information

Center project name: zCi9586

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 158259 bases at least Q40
 Consensus quality: 159771 bases at least Q30
 Consensus quality: 161055 bases at least Q20
 Insert size: 16377; sum-of-contigs
 Quality coverage: 6.48x in Q20 bases; sum-of-contigs Quality
 coverage: 6.96x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 25 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 2920: contig of 2920 bp in length
* 2921 3020: gap of 100 bp
* 3021 8809: contig of 5789 bp in length
* 8810 8909: gap of 100 bp
* 8910 15686: contig of 6777 bp in length
* 15687 15786: gap of 100 bp
* 15787 25724: contig of 9938 bp in length
* 25725 25824: gap of 100 bp
* 25825 48975: contig of 23151 bp in length
* 48976 49075: gap of 100 bp
* 49076 55848: contig of 6773 bp in length
* 55849 55948: gap of 100 bp
* 55949 58553: contig of 2605 bp in length
* 58554 58653: gap of 100 bp
* 58654 63577: contig of 4924 bp in length
* 63578 63677: gap of 100 bp
* 63678 65772: contig of 2095 bp in length
* 65773 65872: gap of 100 bp
* 65873 72740: contig of 6868 bp in length
* 72741 72840: gap of 100 bp
* 72841 76276: contig of 3436 bp in length
* 76277 76376: gap of 100 bp
* 76377 79359: contig of 2983 bp in length
* 79360 79459: gap of 100 bp
* 79460 84912: contig of 5453 bp in length
* 84913 85012: gap of 100 bp
* 85013 90910: contig of 5898 bp in length
* 90911 91010: gap of 100 bp
* 91011 93849: contig of 2839 bp in length
* 93850 93949: gap of 100 bp
* 93950 100061: contig of 6112 bp in length
* 100062 100161: gap of 100 bp
* 100162 103297: contig of 3136 bp in length
* 103298 103397: gap of 100 bp
* 103398 106346: contig of 2949 bp in length
* 106347 106446: gap of 100 bp
* 106447 110847: contig of 4401 bp in length
* 110848 110947: gap of 100 bp
* 110948 126788: contig of 15841 bp in length
* 126789 126888: gap of 100 bp
* 126889 149155: contig of 22267 bp in length
* 149156 149255: gap of 100 bp
* 149256 151944: contig of 2689 bp in length
* 151945 152044: gap of 100 bp
* 152045 155271: contig of 3227 bp in length
* 155272 155372: gap of 100 bp
* 155373 162129: contig of 6758 bp in length
* 162130 162230: gap of 100 bp
* 162231 166117: contig of 3888 bp in length.

```

FEATURES

source

```

1. .166117
   /organism="Danio rerio"
   /mol_type="genomic DNA"
   /db_xref="taxon:7955"
   /clone_lib="CH211-19586"
   /clone_lib="CHOR1-211"
1. .2920

```

misc_feature

```

misc_feature
/feature="assembly fragment:00162"
fragment_chain:1"
3021..8809
/feature="assembly fragment:00807"
fragment_chain:1"
8910..15686
/feature="assembly fragment:01042"
fragment_chain:1"
15787..25724
/feature="assembly fragment:01238"
fragment_chain:1"
25825..48975
/feature="assembly fragment:01616"
fragment_chain:1"
49076..55848
/feature="assembly fragment:00679"
fragment_chain:1"
55949..58553
/feature="assembly fragment:00180"
fragment_chain:2"
58654..63577
/feature="assembly fragment:00363"
fragment_chain:2"
63678..65772
/feature="assembly fragment:00055"
fragment_chain:2"
65873..72740
/feature="assembly fragment:01139"
fragment_chain:2"
72841..76276
/feature="assembly fragment:00563"
fragment_chain:2"
76377..79359
/feature="assembly fragment:00241"
fragment_chain:3"
79460..84912
/feature="assembly fragment:00621"
fragment_chain:3"
85013..90910
/feature="assembly fragment:00877"
fragment_chain:3"
91011..93849
/feature="assembly fragment:00297"
fragment_chain:3"
93950..100061
/feature="assembly fragment:00957"
fragment_chain:4"
100162..103297
/feature="assembly fragment:00268"
fragment_chain:4"
103398..106346
/feature="assembly fragment:00330"
fragment_chain:4"
106447..110847
/feature="assembly fragment:00456"
fragment_chain:4"
110948..126788
/feature="assembly fragment:01401"
fragment_chain:5"
126889..149155
/feature="assembly fragment:02015"
fragment_chain:5"
149256..151944
/feature="assembly fragment:00145"
fragment_chain:5"
152045..155271
/feature="assembly fragment:00405"
fragment_chain:5"
155372..162129
/feature="assembly fragment:00742"
fragment_chain:5"
162230..166117
/feature="assembly fragment:00507"
fragment_chain:5"
clone_end:T7
vector_side:right"

```

ORIGIN

Alignment Scores:
 Pred. No.: 1.3e+04 Length: 166117
 Score: 38.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.0% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-10 (1-9) x CR855265 (1-166117)

Qy 1 PheLeuGlyIleValLeuAlaLeuIle 9

Db 27613 TTCCTTGGATTGTTTTCAGCTTGCTT 27639

RESULT 48

AC114043/c

LOCUS Rattus norvegicus clone CH230-187H10, WORKING DRAFT SEQUENCE. HTG 13-MAY-2003

DEFINITION AC114043

ACCESSION AC114043.4 GI:30581613

VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Rattus;

1 (bases 1 to 213305)

Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angulano,D.,

Anylebechi,V., Aoyagi,A., Ayodeji,M., Baca,B., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barbet,M., Barnstead,M., Benahmed,P.,

Biswalo,K., Blair,J., Blankenburg,K., Blych,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Escoto-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregorgis,S., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,

Gunnaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,

Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

Kowitz,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,

Lorensuhsu,L., Loulseg,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,

Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwankwelen,O., Okwuonu,G., Olarunpesecon,A., Pal,S., Parks,K.,

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L.,

Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P.,

Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savery,G., Scherer,S., Scott,G., Shateman,S., Shen,H.,

Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,

Steinle,M., Strong,N., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,

Valas,R., Vera,V., Villaseña,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished

2 (bases 1 to 213305)

Worley,K.C.

Direct Submission

Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 213305)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GKM

Center clone name: CH230-187H10

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 197649 bases at least Q40

Consensus quality: 199600 bases at least Q30

Consensus quality: 201081 bases at least Q20

Estimated insert size: 210356; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 213305; contig of 213305 bp in length.

Location/Qualifiers

1. 213305

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-187H10"

1. 1105

/note="wgs end extension

/clone_end:Sp6"

3808. 5310

/note="wgs end extension

/clone_end:Sp6"

5361. 7155

FEATURES

source

misc_feature

misc_feature

misc_feature

* be preserved.

* 1 230370: contig of 230370 bp in length
 * 230371 230470: gap of unknown length
 * 230471 232746: contig of 2276 bp in length.

FEATURES

source
 1. 232746
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-6D9"
 1. 1261
 /note="wgs contig"
 89782. 92268
 /note="wgs contig"
 218805 221979
 /note="wgs contig"
 230371. 230470
 /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 1.79e+04 Length: 232746
 Score: 38.00 Matches: 8
 Percent Similarity: 100.0% Conservatives: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.0% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-10 (1-9) x AC125778 (1-232746)

QY 1 PhleuGlyleValleuAlaLeuile 9

DB 209083 TTCTGGGATAGTGTAGCTCTTC 209057

RESULT 50

AC097673/c
 AC097673 271932 bp DNA linear HTG 03-OCT-2002
 Rattus norvegicus clone CH230-64D3, *** SEQUENCING IN PROGRESS ***.

DEFINITION

AC097673

AC097673.5 GI:22855493

HTG, HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 271932)

REFERENCE
 AUTHORS
 Muzny, D., Maric, M., Metzker, M., Lee, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenship, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregeorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hui, Y., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowls, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenzen, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwankweme, O., Okwuonu, G., Olarpunsacon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfanckoch, C.,
 Plummer, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Woodden, H., Yoon, L., Yoon, V.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weinstock, G., and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 271932)
 Morley, K. C.
 Direct Submission
 Submitted (21-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 271932)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (03-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 14, 2002 this sequence version replaced gi:21735993.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
 sequence may extend beyond the ends of the clone and there may be
 contigs that consist entirely of whole genome shotgun sequence
 reads. Both end sequences and whole genome shotgun sequence only
 contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRRR
 Center clone name: CH230-64D3
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 194955 bases at least Q40
 Consensus quality: 199813 bases at least Q30
 Consensus quality: 202986 bases at least Q20
 Estimated insert size: 220560; sum-of-contigs estimation
 Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces
 is believed to be correct as given, however the sizes
 of the gaps between them are based on estimates that have
 been provided by the submitter.
 * This sequence will be replaced
 * By the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 271932: contig of 271932 bp in length.


```
FEATURES
  source
    Location/Qualifiers
      1. .271932
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-64D3"
      misc_feature
        1. .1087
          /note="wgs contig"
      misc_feature
        101935..103160
          /note="wgs contig"
      misc_feature
        238240..239483
          /note="wgs contig"
      misc_feature
        246984..248761
          /note="wgs contig"
      misc_feature
        270876..271932
          /note="wgs contig"
  ORIGIN
    Alignment Scores:
      Pred. No.:      2.07e+04      Length:      271932
      Score:         38.00          Matches:      8
      Percent Similarity: 100.0%      Conservative: 1
      Best Local Similarity: 88.9%      Mismatches:  0
      Query Match:     95.0%          Indels:       0
      DB:              14             Gaps:         0
    US-10-774-176-10 (1-9) x AC097673 (1-271932)
    Oy      1 PheLeuGlyIleValLeuAlaLeuile 9
      |||||||
    Db      165262 TTCTTGGGATAGTGTAGCTCTTC 165236
```

Search completed: April 25, 2006, 20:36:46
Job time : 3098.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds

(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-9

Perfect score: 48

Sequence: 1 GLPHIRVFL 9

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/abss/ABSSWEB spool/US10774176/runat 24042006 165112 19185/app query.fasta.1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT5=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOALIGN=200 -THR_SCORE=spect -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05p
-USER=US10774176 @CGEN 1 1 3463 @runat 24042006 165112 19185 -NCFU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 21.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	130	10	ADK11793
2	48	100.0	299	10	ACD93536
3	48	100.0	453	5	AAS87174
4	48	100.0	475	13	ADU11677

5	48	100.0	927	6	ABT07721
6	48	100.0	927	8	ABX76333
7	48	100.0	927	10	ADB80503
8	48	100.0	927	11	ADN38723
9	48	100.0	973	8	AAD56198
10	48	100.0	1156	6	ABV99349
11	48	100.0	1263	3	AAA27058
12	48	100.0	1331	8	AAD56199
13	48	100.0	2020	10	ADJ56299
14	48	100.0	2053	8	ACC51052
15	48	100.0	2053	8	ABX76332
16	48	100.0	2053	8	AAD56197
17	48	100.0	2053	8	AAD56200
18	48	100.0	2053	11	ADN38721
19	48	100.0	2053	12	ADL06473
20	48	100.0	2053	12	ADN03961
21	48	100.0	2053	13	ADR25444
22	48	100.0	2053	13	ACN38510
23	48	100.0	2053	13	ADV35098
24	48	100.0	2338	5	AAS87175
25	48	100.0	2359	4	AAK94253
26	48	100.0	2359	12	ADL30831
27	48	100.0	2361	4	AAK94254
28	48	100.0	2361	12	ADI26162
29	48	100.0	2361	12	ADL30833
30	42	87.5	1587	13	ADS96565
31	42	87.5	1704	4	ABL20593
32	42	87.5	4357	4	ABL20592
33	41	85.4	901	3	ACN80343
34	41	85.4	1260	6	AAA27060
35	41	85.4	1260	3	ABK87175
36	41	85.4	1260	10	ADB97513
37	41	85.4	1263	4	AAF89736
38	41	85.4	1263	6	AK87174
39	41	85.4	110000	12	ADO34435_3
40	41	85.4	110000	12	ADO34435_4
41	41	85.4	110000	12	ADO34435_4
42	40	83.3	416	8	ABX36840
43	40	83.3	433	8	ABX49697
44	40	83.3	465	9	ACH45478
45	40	83.3	467	9	ACH45478
46	40	83.3	515	12	ACH74590
47	40	83.3	565	5	ABV58745
48	40	83.3	615	5	ABV56637
49	40	83.3	933	5	AAH52150
50	40	83.3	1333	12	ADQ87043
51	40	83.3	1349	12	ADQ87041
52	40	83.3	1357	4	AAI60576
53	40	83.3	1364	14	ADM03242
54	40	83.3	1370	4	AAI58790
55	40	83.3	1370	5	ADQ99010
56	40	83.3	1370	9	ADB48770
57	40	83.3	1377	3	AACT75984
58	40	83.3	1568	6	ABQ54666
59	40	83.3	2374	10	ADA52605
60	40	83.3	3613	4	AAK80972
61	40	83.3	3621	4	AAK80975
62	40	83.3	3623	4	AAK80973
63	40	83.3	110000	6	ABA90521_10
64	39	81.2	385	8	AAH42094
65	39	81.2	1884	4	AAH41367
66	39	81.2	1884	6	ABK62585
67	39	81.2	1884	10	ADB56207
68	39	81.2	1884	10	ADB53152
69	39	81.2	1884	11	ADM22417
70	39	81.2	1884	11	ADM22314
71	39	81.2	2019	13	ADRO7253
72	39	81.2	2276	13	ADO82034
73	39	81.2	2977	12	ADQ63251
74	39	81.2	3498	12	ADI42189
75	39	81.2	3498	12	ADO02700
76	39	81.2	3498	12	ADO62321
77	39	81.2	18997	6	AAAD36072

Adk11793 Breast ca
Acd93536 Human col
Aas87174 DNA encod
Adu11677 Solid tum

c 78	39	81.2	87394	13	ADT55151	Adt55151 Nucleotid	c 151	37	77.1	1087	11	ACN84111	Acn84111 Breast ca
c 79	39	81.2	110000	6	ABA03041.28	Continuation (29 o	c 152	37	77.1	1110	3	AAC45795	Aac45795 Arabidops
c 80	39	81.2	181257	12	ADF69677	Adf69677 Human SLC	c 153	37	77.1	1197	3	ADU05454	Adu05454 DNA encod
c 81	38	79.2	357	2	AX221108	Aax221108 Polynucle	c 154	37	77.1	1200	4	AAS53681	Aas53681 Helicobac
c 82	38	79.2	425	14	ABE65613	AbE65613 Rice geno	c 155	37	77.1	1200	8	ACA34860	Aca34860 Prokaryot
c 83	38	79.2	601	8	ABX61872	Abx61872 Novel hum	c 156	37	77.1	1380	3	ACA37623	Aca37623 Arabidops
c 84	38	79.2	648	4	AAH32164	Aah32164 Human olf	c 157	37	77.1	1429	2	AAV04237	Aav04237 Arabidops
c 85	38	79.2	663	13	ADH42907	Adt42907 Bacterial	c 158	37	77.1	1448	2	AAV07964	Aav07964 Helicobac
c 86	38	79.2	933	4	AAH32118	Aah32118 Human olf	c 159	37	77.1	1516	4	AAS56970	Aas56970 C. tracho
c 87	38	79.2	936	5	AAS42443	Aas42443 Human CDN	c 160	37	77.1	1516	10	ADD42764	Add42764 Chlamydia
c 88	38	79.2	936	6	ABK16621	Abk16621 Human G-c	c 161	37	77.1	1517	6	ABQ76468	Abq76468 S. cerevi
c 89	38	79.2	936	6	ABZ43127	Abz43127 Human GPC	c 162	37	77.1	1533	11	ACH99608	Ach99608 Klebellei
c 90	38	79.2	936	6	ABK68450	Abk68450 Human DNA	c 163	37	77.1	1620	11	ACL26343	Acl26343 Rice abio
c 91	38	79.2	936	6	ABK37729	Abk37729 DNA encod	c 164	37	77.1	1620	12	ADI45106	Adi45106 Rice isop
c 92	38	79.2	936	10	ABZ77872	Abz77872 Human G P	c 165	37	77.1	1629	8	ABT19170	Abt19170 Aspergill
c 93	38	79.2	940	6	ABK33474	Abk33474 Human CDN	c 166	37	77.1	1680	6	ABK75035	Abk75035 Bacillus
c 94	38	79.2	1021	6	ABK33473	Abk33473 Human CDN	c 167	37	77.1	1686	10	ADG75807	Adg75807 Human pro
c 95	38	79.2	1021	6	ABK33472	Abk33472 Human CDN	c 168	37	77.1	1813	4	AAH18471	Aah18471 Human cDN
c 96	38	79.2	1021	11	ADM29635	Adm29635 Novel hum	c 169	37	77.1	1813	8	ABT18576	Abt18576 Aspergill
c 97	38	79.2	1021	11	ADM29633	Adm29633 Novel hum	c 170	37	77.1	1814	5	AAF68401	Aaf68401 Human lun
c 98	38	79.2	1101	10	ADF00145	Adf00145 Bacterial	c 171	37	77.1	1814	6	ABK38312	Abk38312 CDNA enco
c 99	38	79.2	1336	10	ADC85686	Adc85686 Human GPC	c 172	37	77.1	1814	8	ACA10641	Aca10641 Human lun
c 100	38	79.2	1935	8	ACA21202	Aca21202 Prokaryot	c 173	37	77.1	1814	8	ABX99592	Abx99592 Lung can
c 101	38	79.2	1950	10	ADC92551	Adc92551 E. faeciu	c 174	37	77.1	1814	12	ADH45838	Adh45838 Human lun
c 102	38	79.2	2055	9	ADA30368	Ada30368 DNA encod	c 175	37	77.1	1814	12	ADE72375	AdE72375 Human lun
c 103	38	79.2	2309	4	ABL24854	Ab124854 Drosophil	c 176	37	77.1	1814	13	ADJ19757	Adj19757 Human lun
c 104	38	79.2	2613	13	ADQ88145	Adq88145 Zea mays	c 177	37	77.1	1815	6	ASQ88162	Asq88162 Human ost
c 105	38	79.2	2796	8	ACA33757	Aca33757 Prokaryot	c 178	37	77.1	1815	6	ABN95798	Abn95798 Gene #229
c 106	38	79.2	2884	13	ADQ88143	Adq88143 Zea mays	c 179	37	77.1	1815	10	ADF81422	Adf81422 Leukaemia
c 107	38	79.2	3029	10	ADA52973	Ada52973 Human cod	c 180	37	77.1	1896	8	ABT20990	Abt20990 Aspergill
c 108	38	79.2	3068	13	ADX63820	Adx63820 Plant ful	c 181	37	77.1	1938	12	ADQ83372	Adq83372 Human tum
c 109	38	79.2	3126	12	ADJ42190	Adi42190 Plant tra	c 182	37	77.1	1988	10	ADE31346	AdE31346 Human dia
c 110	38	79.2	3126	12	ADO02701	Ado02701 Corn orth	c 183	37	77.1	1997	8	ACC46113	Acc46113 Human dit
c 111	38	79.2	3126	12	ADO62322	Ado62322 Transcrip	c 184	37	77.1	2013	12	ADP28590	Adp28590 Human sec
c 112	38	79.2	3351	13	ADS47394	Ads47394 Bacterial	c 185	37	77.1	2018	6	ABQ99530	Abq99530 Human cod
c 113	38	79.2	5080	4	ABL09204	Ab109204 Drosophil	c 186	37	77.1	2044	4	AAS57015	Aas57015 C. tracho
c 114	38	79.2	28564	10	ADA47107	Ada47107 Rat gene	c 187	37	77.1	2044	10	ADD42809	Add42809 Chlamydia
c 115	38	79.2	28564	10	ADA47113	Ada47113 Rat gene	c 188	37	77.1	2044	13	ADS60516	Ads60516 Bacterial
c 116	38	79.2	28564	10	ADA47117	Ada47117 Rat gene	c 189	37	77.1	2099	8	ABT20392	Abt20392 Aspergill
c 117	38	79.2	28564	10	ADA47111	Ada47111 Rat gene	c 190	37	77.1	2162	8	ACC44310	Acc44310 Gene enco
c 118	38	79.2	28564	10	ADH83334	Adh83334 Rat gene	c 191	37	77.1	2198	8	ABL25468	Ab125468 Drosophil
c 119	38	79.2	28564	10	ADE57550	AdE57550 Rat gene	c 192	37	77.1	2463	8	ACA57533	Aca57533 Human adi
c 120	38	79.2	28564	10	ADE57554	AdE57554 Rat gene	c 193	37	77.1	2540	4	ABL24550	Ab124550 Drosophil
c 121	38	79.2	37515	6	ABQ66998	Abq66998 Human ang	c 194	37	77.1	2689	10	ADB62112	Adb62112 Human cDN
c 122	38	79.2	40645	9	ADA49717	Ada49717 Genomic D	c 195	37	77.1	2736	6	ABA04268	AbA04268 Human ATP
c 123	38	79.2	40645	9	ADA49717	Ada49717 Genomic D	c 196	37	77.1	2864	3	AAF18312	Aaf18312 Lung can
c 124	38	79.2	65047	11	ACN44020	Acn44020 Mouse gen	c 197	37	77.1	3159	4	AAS57035	Aas57035 C. tracho
c 125	38	79.2	93323	12	ADQ97047_3	Continuation (4 of	c 198	37	77.1	3159	6	ABL92695	Ab192695 Chlamydia
c 126	38	79.2	134525	2	AAQ04525	Aaq04525 Total bas	c 199	37	77.1	3159	10	ADD42669	Add42669 Chlamydia
c 127	38	79.2	231222	10	ADL33693	Adl33693 Osteoarth	c 200	37	77.1	3196	13	ADR08166	Adr08166 Full leng
c 128	37	77.1	246	12	ADP91011	Adp91011 Cotton ex	c 201	37	77.1	3215	12	ADQ85333	Adq85333 Human tum
c 129	37	77.1	328	4	AAS38693	Aas38693 Novel hum	c 202	37	77.1	3237	12	ADJ62801	Adj62801 Human cDN
c 130	37	77.1	445	6	ABT06987	Abt06987 Human ova	c 203	37	77.1	3237	14	ADZ26147	Adz26147 Human gen
c 131	37	77.1	445	8	ABX72865	Abx72865 Human ova	c 204	37	77.1	3244	8	ACD13424	AcD13424 Human DNA
c 132	37	77.1	484	4	AA192734	Aa192734 Human pol	c 205	37	77.1	3535	12	ADQ67517	Adq67517 Novel hum
c 133	37	77.1	497	9	ACH34124	Ach34124 Human end	c 206	37	77.1	3618	12	ADM67140	Adm67140 Murine ad
c 134	37	77.1	514	5	AAS73460	Abx73460 DNA encod	c 207	37	77.1	3791	6	ABK52289	Abk52289 CDNA enco
c 135	37	77.1	553	14	ACL55062	ACL55062 Human col	c 208	37	77.1	3849	2	AAV12271	Aav12271 Temperate
c 136	37	77.1	652	6	ABX66890	Abx66890 Helicobac	c 209	37	77.1	4054	4	ABL13096	Ab113096 Drosophil
c 137	37	77.1	662	6	ABX66333	Abx66333 Helicobac	c 210	37	77.1	4097	10	ADA53501	Ada53501 Human cod
c 138	37	77.1	708	12	ADQ22331	Adq22331 Human sof	c 211	37	77.1	4099	8	ABT19796	Abt19796 Aspergill
c 139	37	77.1	732	4	AA114054	Aa114054 Human bre	c 212	37	77.1	4506	8	ABL05213	Ab105213 Drosophil
c 140	37	77.1	806	3	AAAC63429	Aac63429 Human sec	c 213	37	77.1	4506	10	ADB69186	Adb69186 C. neofo
c 141	37	77.1	806	8	ABZ73602	Abz73602 Secreted	c 214	37	77.1	4880	10	ADBE62039	Adbe62039 Rat gene
c 142	37	77.1	806	8	ADA98104	Ada98104 Human sec	c 215	37	77.1	6121	10	ADBE62035	Adbe62035 Rat gene
c 143	37	77.1	806	10	ABZ67196	Abz67196 Human sec	c 216	37	77.1	6121	10	ADBE62035	Adbe62035 Rat gene
c 144	37	77.1	906	2	AAV37157	Aav37157 DNA sequ	c 217	37	77.1	6588	2	AAV04238	Aav04238 Arabidops
c 145	37	77.1	906	4	AAH01749	Aah01749 Helicobac	c 218	37	77.1	7066	10	ADBE62108	Adbe62108 Rat gene
c 146	37	77.1	917	4	AAH05062	Aah05062 Human cDN	c 219	37	77.1	12276	4	ABL05212	Ab105212 Drosophil
c 147	37	77.1	1014	8	ABZ52580	Abz52580 Aspergill	c 220	37	77.1	18925	6	ABA81553	AbA81553 Human pho
c 148	37	77.1	1017	10	ACC61132	Acc61132 Gene sequ	c 221	37	77.1	18925	6	AAS94692	Aas94692 Human pho
c 149	37	77.1	1017	10	ADK62513	Adk62513 Disease t	c 222	37	77.1	18925	6	AAS94565	Aas94565 Human pho
c 150	37	77.1	1017	13	ADT47657	Adt47657 Bacterial	c 223	37	77.1	21407	4	ABL28892	Ab128892 Drosophil

C 224	37	77.1	27413	11	ACN44832	Acn44832 Mouse gen	C 297	36	75.0	1534	12	ADJ67053	Adj67053 Human sec
C 225	37	77.1	27901	13	ABD33593	Abd33593 Murine ca	C 298	36	75.0	1594	3	AAC59913	Aac59913 Human sec
C 226	37	77.1	27901	14	AD213316	Ad213316 Murine ca	C 299	36	75.0	1616	6	ABQ76427	Abq76427 S. cerevi
C 227	37	77.1	28564	10	AD859178	Ad859178 Human gen	C 300	36	75.0	1641	4	ABL13415	Ab113415 Drosophil
C 228	37	77.1	36351	12	ADQ59173	Adq59173 MSI-H car	C 301	36	75.0	1692	13	ADL12664	Adl12664 Plant ful
C 229	37	77.1	48244	9	ADA02660	Ada02660 Human SEM	C 302	36	75.0	1714	10	ADB58624	Adb58624 Toxigity-
C 230	37	77.1	48244	10	ADB72398	Adb72398 Human SEM	C 303	36	75.0	1714	10	ADB53275	Adb53275 Primary r
C 231	37	77.1	48244	10	AD895908	Ad895908 Human SEM	C 304	36	75.0	1725	13	ADU01647	Adu01647 Novel hum
C 232	37	77.1	49159	14	AB896522	Ab896522 Human NPK	C 305	36	75.0	1740	6	ABK73188	Abk73188 Bacillus
C 233	37	77.1	58985	9	AB259738	Ab259738 Human sec	C 306	36	75.0	1866	6	ABK42194	Abk42194 Genomic s
C 234	37	77.1	77287	9	AAD58261	Aad58261 Murine tu	C 307	36	75.0	1886	9	ADB60350	Adb60350 Connectiv
C 235	37	77.1	77287	9	AAD58261	Aad58261 Murine tu	C 308	36	75.0	1920	4	ABL23977	Ab123977 Drosophil
C 236	37	77.1	86080	6	ABQ88164	Abq88164 Human ost	C 309	36	75.0	1939	13	ADG84288	Adg84288 Human TMD
C 237	37	77.1	86080	6	ABK83561	Abk83561 Human cDN	C 310	36	75.0	1949	10	ADG84288	Adg84288 Human TMD
C 238	37	77.1	86080	10	AD771054	Ad771054 Human pro	C 311	36	75.0	1955	10	ADA53961	Ada53961 Human cod
C 239	37	77.1	86080	12	ADQ18878	Adq18878 Human sof	C 312	36	75.0	2000	11	ACL37864	Ac137864 Rice stre
C 240	37	77.1	96592	9	ADA02612	Ada02612 Human NPK	C 313	36	75.0	2047	5	ACS87283	Aas87283 DNA encod
C 241	37	77.1	96592	10	AD872350	Adh72350 Human NPK	C 314	36	75.0	2074	11	ADN02258	Adm02258 Human cDN
C 242	37	77.1	96592	10	AD872350	Adh72350 Human NPK	C 315	36	75.0	2169	6	ABN71170	Abn71170 Streptoco
C 243	37	77.1	96592	10	AD872350	Adh72350 Human NPK	C 316	36	75.0	2196	6	ABN67732	Abn67732 Streptoco
C 244	37	77.1	96592	10	AD872350	Adh72350 Human NPK	C 317	36	75.0	2199	13	ADV84451	Adv84451 Streptoco
C 245	37	77.1	100137	12	ADQ97643	Adq97643 Human can	C 318	36	75.0	2274	5	ABA19513	Abal19513 Human ner
C 246	37	77.1	110000	2	AAZ01425_05	Continuation (6 of	C 319	36	75.0	2291	6	ADG79245	Adg79245 Human sec
C 247	37	77.1	117829	12	ADQ97319	Adh97319 Human can	C 320	36	75.0	2336	3	AAA47422	Aaa47422 Sequence
C 248	37	77.1	118518	12	ADH93063	Adh93063 Human NF-	C 321	36	75.0	2557	12	ADI26160	Adi26160 Human cDN
C 249	37	77.1	136489	14	AD212560	Ad212560 Human can	C 322	36	75.0	2557	12	ADI26160	Adi26160 Human cDN
C 250	37	77.1	143601	8	AB222654	Ab222654 Human epi	C 323	36	75.0	2557	12	ADI26160	Adi26160 Human cDN
C 251	37	77.1	167739	9	AAD58258	Aad58258 Murine tu	C 324	36	75.0	2557	12	ADI26160	Adi26160 Human cDN
C 252	37	77.1	167739	9	AAD58258	Aad58258 Murine tu	C 325	36	75.0	3051	4	ABL25570	Ab125570 Drosophil
C 253	37	77.1	186510	10	AD824797	Ad824797 Human end	C 326	36	75.0	3053	10	ADB62545	Adb62545 Human cDN
C 254	37	77.1	192639	10	ADL13676	Adl13676 Osteoarth	C 327	36	75.0	3148	8	ACA46749	Acc46749 Human dit
C 255	36	75.0	112	5	ABA19514	Abal19514 Human gen	C 328	36	75.0	3240	6	ABA92243	Ab92243 Corynebac
C 256	36	75.0	220	9	ACH4867	Ach4867 Human fce	C 329	36	75.0	3341	4	ABL13414	Ab113414 Drosophil
C 257	36	75.0	284	12	ADH00285	Adh00285 Kidney di	C 330	36	75.0	4210	13	ADRO8156	Adro8156 Full leng
C 258	36	75.0	299	3	AAA87482	Aaa87482 Rat hepat	C 331	36	75.0	4468	10	ADG62116	Adg62116 Rat gene
C 259	36	75.0	339	5	ABA13435	Abal13435 Human ner	C 332	36	75.0	4468	10	ADG62116	Adg62116 Rat gene
C 260	36	75.0	339	12	ADQ34506	Ado34506 Human SLI	C 333	36	75.0	4468	10	ADG62116	Adg62116 Rat gene
C 261	36	75.0	339	12	ADQ34506	Ado34506 Human SLI	C 334	36	75.0	4468	10	ADG62116	Adg62116 Rat gene
C 262	36	75.0	438	13	ADG65674	Adg65674 Cotton CD	C 335	36	75.0	4478	8	ABT17690	Abt17690 Aspergill
C 263	36	75.0	450	12	ACH84105	Ach84105 Human gen	C 336	36	75.0	4498	4	ABL09726	Ab109726 Drosophil
C 264	36	75.0	474	4	ABS32918	Abs32918 Human liv	C 337	36	75.0	4631	3	AAA80620	Aaa80620 Human sec
C 265	36	75.0	474	6	ABS07998	Abs07998 Human gen	C 338	36	75.0	4631	9	ADA27044	Ada27044 Human nov
C 266	36	75.0	476	9	ACH28049	Ach28049 Human adu	C 339	36	75.0	4631	12	AD865574	Ad865574 Novel hum
C 267	36	75.0	481	8	ACC60299	Acc60299 Rice endo	C 340	36	75.0	4682	8	ABT19504	Abt19504 Aspergill
C 268	36	75.0	518	12	ACH70405	Ach70405 Human gen	C 341	36	75.0	4779	8	ACC72077	Acc72077 Human NOV
C 269	36	75.0	524	4	ABA61551	Abac61551 Human fce	C 342	36	75.0	4986	6	ABX04161	Abx04161 Human mRN
C 270	36	75.0	524	4	AA141464	Aai41464 Probe #10	C 343	36	75.0	4986	12	ADQ19967	Adq19967 Human sof
C 271	36	75.0	524	4	AAK35748	Aak35748 Human bon	C 344	36	75.0	4995	3	AAA80612	Aaa80612 Human ITG
C 272	36	75.0	524	4	AAK09854	Aak09854 Human bra	C 345	36	75.0	4995	8	ACC72076	Acc72076 Human NOV
C 273	36	75.0	549	13	ADQ48714	Adq48714 Novel can	C 346	36	75.0	4995	9	ADA27036	Ada27036 Human nov
C 274	36	75.0	575	13	ADQ48714	Adq48714 Plant ful	C 347	36	75.0	4995	12	AD865566	Ad865566 Novel hum
C 275	36	75.0	673	3	ABA97721	Ab97721 Pepper mi	C 348	36	75.0	5042	5	AAC91901	Aac91901 Human A25
C 276	36	75.0	676	10	ADD15958	Add15958 cDNA (Seq	C 349	36	75.0	5042	5	AAS16873	Aas16873 Human A25
C 277	36	75.0	760	13	ADX12177	Adx12177 Plant ful	C 350	36	75.0	5117	12	ADQ24050	Adq24050 Human sof
C 278	36	75.0	788	13	ADQ55235	Adq55235 Novel can	C 351	36	75.0	5269	8	ABZ09866	Abz09866 Human 5'
C 279	36	75.0	805	13	ADX47203	Adx47203 Plant ful	C 352	36	75.0	5427	7	ADR41320	Adr41320 Human CD-
C 280	36	75.0	900	5	ACS05625	Aas05625 Mammalian	C 353	36	75.0	5569	4	ABL10364	Ab110364 Drosophil
C 281	36	75.0	938	8	ACC46182	Acc46182 Human dit	C 354	36	75.0	5599	10	AD807773	Ad807773 Novel cod
C 282	36	75.0	940	11	ADM29637	Adm29637 Novel hum	C 355	36	75.0	6222	6	ADQ34176	Adq34176 Fugu rubr
C 283	36	75.0	960	4	ABL25571	Ab125571 Drosophil	C 356	36	75.0	6229	14	ADZ45055	Adz45055 Plasmid p
C 284	36	75.0	962	4	AAH32281	Aah32281 Human olf	C 357	36	75.0	6921	13	AD897706	Ad897706 Rabbit be
C 285	36	75.0	970	4	AAH31589	Aah31589 Human olf	C 358	36	75.0	6973	4	ABL23976	Ab123976 Drosophil
C 286	36	75.0	1001	12	ACH89835	Ach89835 Human gen	C 359	36	75.0	7397	2	AAV60578	Aav60578 Human tum
C 287	36	75.0	1135	4	ABK42193	Abk42193 Genomic s	C 360	36	75.0	7422	13	ADT47466	Adt47466 Bacterial
C 288	36	75.0	1135	9	ADB60349	Adb60349 Connectiv	C 361	36	75.0	7862	4	ABL07766	Ab107766 Drosophil
C 289	36	75.0	1236	11	ABD01038	Abd01038 Klebsiell	C 362	36	75.0	8169	13	AD888996	Ad888996 Human ARH
C 290	36	75.0	1281	3	AAA27059	Aaa27059 Mouse 574	C 363	36	75.0	11541	12	AD030133	Ad030133 Mouse GPC
C 291	36	75.0	1311	10	ADC87254	Adc87254 Human GPC	C 364	36	75.0	12801	13	AD897709	Ad897709 Rabbit al
C 292	36	75.0	1358	10	ADC08037	Adc08037 Rice DNA	C 365	36	75.0	13161	5	AA97867	Aaf97867 Human neu
C 293	36	75.0	1388	10	AD858528	Ad858528 Human gen	C 366	36	75.0	14910	5	AA84004	Aas84004 DNA encod
C 294	36	75.0	1388	10	AD447423	Ad447423 Human gen	C 367	36	75.0	15932	8	ABZ73902	Abz73902 Secreted
C 295	36	75.0	1446	12	ADN98947	Adn98947 Novel hum	C 368	36	75.0	15932	8	ADA44289	Ada44289 Human sec
C 296	36	75.0	1446	12	ADN00516	Adn00516 Novel hum	C 369	36	75.0	19974	6	AB865025	Ab865025 Invertebr

370	36	75.0	20978	4	ABL20786	Abi20786 Drosophil	c 443	35	72.9	593	4	AAK11976	Aak11976 Human bra
371	36	75.0	31737	10	ACF67733	Acf67733 Nucleohab	c 444	35	72.9	593	4	ABS37332	Abs37332 Human liv
372	36	75.0	43226	2	AX60263	Aax60263 Photoic a	c 445	35	72.9	593	6	ABS11670	Abs11670 Human gen
373	36	75.0	43539	13	ADV87729	Adv87729 Streptoco	446	35	72.9	596	13	ADO55643	Ado55643 Novel can
374	36	75.0	43539	13	ADV78982	Adv78982 Streptoco	446	35	72.9	605	6	ABQ66346	Abq66346 Arabidops
375	36	75.0	50000	9	ADB16929	Adb16929 Human DYX	c 448	35	72.9	613	4	ABL04135	Abi04135 Drosophil
376	36	75.0	51664	11	ACN44432	Acn44432 Mouse gen	c 449	35	72.9	662	11	ADM92651	Adm92651 SNP-conta
377	36	75.0	72314	11	ACN44494	Acn44494 Mouse gen	450	35	72.9	679	6	ABQ58871	Abq58871 Human col
378	36	75.0	76826	14	ABE33172_3	Continuation (4 of	451	35	72.9	786	11	ACH98291	Ach98291 Klebaieil
379	36	75.0	90141	12	ADQ97867	Adq97867 Mouse can	452	35	72.9	825	8	ACA35019	ACA35019 Prokaryot
380	36	75.0	92112	13	ADQ99457_3	Continuation (4 of	453	35	72.9	843	13	ADA46238	Ada46238 Bacteriat
381	36	75.0	98345	13	ABD32892	Abd32892 Human can	454	35	72.9	855	4	AAH70354	Aah70354 Human cer
382	36	75.0	107304	13	ABD33230	Abd33230 Murine ca	c 455	35	72.9	858	3	AAH70354	AAH70354 Human cer
383	36	75.0	110000	6	ABN71527_20	Continuation (21 o	c 456	35	72.9	858	3	AAH70354	AAH70354 Human cer
384	36	75.0	110000	10	ADG70184	Adg70184 DNA of BA	c 457	35	72.9	858	3	AAH70354	AAH70354 Human cer
385	36	75.0	110000	10	ACF65385_2	Continuation (3 of	458	35	72.9	858	3	AAH70354	AAH70354 Human cer
386	36	75.0	110000	10	ACF67367_03	Continuation (4 of	459	35	72.9	858	3	AAH70354	AAH70354 Human cer
387	36	75.0	110000	10	ACF42745_0	ACF42745 Human ALM	460	35	72.9	858	3	AAH70354	AAH70354 Human cer
388	36	75.0	110000	13	ADQ32921_6	Continuation (7 of	461	35	72.9	858	3	AAH70354	AAH70354 Human cer
389	36	75.0	110000	13	ADV81204_21	Continuation (22 o	c 462	35	72.9	858	3	AAH70354	AAH70354 Human cer
390	36	75.0	110000	14	ABE42737_09	Continuation (10 o	463	35	72.9	858	3	AAH70354	AAH70354 Human cer
391	36	75.0	124856	12	ABD333616	Abd333616 Human can	464	35	72.9	858	3	AAH70354	AAH70354 Human cer
392	36	75.0	137000	12	ADH77370	Adh77370 Human PTP	465	35	72.9	858	3	AAH70354	AAH70354 Human cer
393	36	75.0	150351	13	ABD33360	Abd33360 Murine ca	c 466	35	72.9	858	3	AAH70354	AAH70354 Human cer
394	36	75.0	154902	6	ABQ88198	Abq88198 Human oet	c 467	35	72.9	858	3	AAH70354	AAH70354 Human cer
395	36	75.0	215126	11	ADQ97362	Adq97362 Mouse can	c 468	35	72.9	858	3	AAH70354	AAH70354 Human cer
396	36	75.0	254493	12	ACN44514	Acn44514 Human gen	c 469	35	72.9	858	3	AAH70354	AAH70354 Human cer
397	36	75.0	265118	5	AHA41227	Aha41227 Pyrococcu	c 470	35	72.9	858	3	AAH70354	AAH70354 Human cer
398	36	75.0	321491	11	ACN44202	Acn44202 Human gen	471	35	72.9	858	3	AAH70354	AAH70354 Human cer
399	36	75.0	349980	5	AAH68533	Aah68533 C glutami	472	35	72.9	858	3	AAH70354	AAH70354 Human cer
400	35.5	74.0	643	13	ADO57624	Ado57624 Novel can	473	35	72.9	858	3	AAH70354	AAH70354 Human cer
401	35	72.9	32	2	AAQ86648	Aaq86648 PTP spik	474	35	72.9	858	3	AAH70354	AAH70354 Human cer
402	35	72.9	50	6	ABZ00267	Abz00267 Human leu	475	35	72.9	858	3	AAH70354	AAH70354 Human cer
403	35	72.9	102	12	ACH84049	Ach84049 Human oet	476	35	72.9	858	3	AAH70354	AAH70354 Human cer
404	35	72.9	199	3	AAA44082	Aaa44082 Human sec	477	35	72.9	858	3	AAH70354	AAH70354 Human cer
405	35	72.9	204	4	AAI00785	Aai00785 Human rep	478	35	72.9	858	3	AAH70354	AAH70354 Human cer
406	35	72.9	218	5	ABV58221	Abv58221 Human pro	479	35	72.9	858	3	AAH70354	AAH70354 Human cer
407	35	72.9	269	12	ADP64250	Adp64250 Maize car	480	35	72.9	858	3	AAH70354	AAH70354 Human cer
408	35	72.9	280	12	ADP64251	Adp64251 Maize car	c 481	35	72.9	858	3	AAH70354	AAH70354 Human cer
409	35	72.9	292	12	ADP63244	Adp63244 Maize car	482	35	72.9	858	3	AAH70354	AAH70354 Human cer
410	35	72.9	305	10	ABX87277	Abx87277 Corn ear-	483	35	72.9	858	3	AAH70354	AAH70354 Human cer
411	35	72.9	318	6	ABK46155	Abk46155 CDNA enco	484	35	72.9	858	3	AAH70354	AAH70354 Human cer
412	35	72.9	319	6	ABK46154	Abk46154 CDNA enco	485	35	72.9	858	3	AAH70354	AAH70354 Human cer
413	35	72.9	320	6	ABV94288	Abv94288 Breast ca	c 486	35	72.9	858	3	AAH70354	AAH70354 Human cer
414	35	72.9	354	12	ADJ39141	Adj39141 Plant CDN	c 487	35	72.9	858	3	AAH70354	AAH70354 Human cer
415	35	72.9	356	3	AAO36399	Aao36399 Human sec	c 488	35	72.9	858	3	AAH70354	AAH70354 Human cer
416	35	72.9	359	12	ADN13716	Adn13716 Human pro	489	35	72.9	858	3	AAH70354	AAH70354 Human cer
417	35	72.9	372	12	ADP63257	Adp63257 Maize car	490	35	72.9	858	3	AAH70354	AAH70354 Human cer
418	35	72.9	377	2	AAQ25540	Aaq25540 S gene of	c 491	35	72.9	858	3	AAH70354	AAH70354 Human cer
419	35	72.9	377	2	AAQ25540	Aaq25540 S gene of	492	35	72.9	858	3	AAH70354	AAH70354 Human cer
420	35	72.9	382	6	ABN76153	Abn76153 Human ORF	493	35	72.9	858	3	AAH70354	AAH70354 Human cer
421	35	72.9	395	12	ADQ04817	Adq04817 Maize leu	494	35	72.9	858	3	AAH70354	AAH70354 Human cer
422	35	72.9	399	5	ABV37256	Abv37256 Human pro	495	35	72.9	858	3	AAH70354	AAH70354 Human cer
423	35	72.9	405	5	ABV16252	Abv16252 Human pro	496	35	72.9	858	3	AAH70354	AAH70354 Human cer
424	35	72.9	421	12	ADP64260	Adp64260 Maize car	497	35	72.9	858	3	AAH70354	AAH70354 Human cer
425	35	72.9	425	6	ABK46126	Abk46126 CDNA enco	498	35	72.9	858	3	AAH70354	AAH70354 Human cer
426	35	72.9	427	13	ADU12229	Adu12229 Solid tum	499	35	72.9	858	3	AAH70354	AAH70354 Human cer
427	35	72.9	432	8	ABX50624	Abx50624 Bovine ES	500	35	72.9	858	3	AAH70354	AAH70354 Human cer
428	35	72.9	469	5	ABV46050	Abv46050 Human pro	501	35	72.9	858	3	AAH70354	AAH70354 Human cer
429	35	72.9	482	9	ACH15222	Ach15222 Human adu	502	35	72.9	858	3	AAH70354	AAH70354 Human cer
430	35	72.9	484	9	ACH43211	Ach43211 Human foe	503	35	72.9	858	3	AAH70354	AAH70354 Human cer
431	35	72.9	537	9	ACD10971	Acd10971 Human col	504	35	72.9	858	3	AAH70354	AAH70354 Human cer
432	35	72.9	549	4	ADL04205	Adl04205 Human rep	505	35	72.9	858	3	AAH70354	AAH70354 Human cer
433	35	72.9	557	12	ACH70349	Ach70349 Human gen	506	35	72.9	858	3	AAH70354	AAH70354 Human cer
434	35	72.9	573	2	AAQ25533	Aaq25533 DNA enco	507	35	72.9	858	3	AAH70354	AAH70354 Human cer
435	35	72.9	584	5	ABK18888	Abk18888 Human ner	c 508	35	72.9	858	3	AAH70354	AAH70354 Human cer
436	35	72.9	585	11	ABD13846	Abd13846 Pseudomon	c 509	35	72.9	858	3	AAH70354	AAH70354 Human cer
437	35	72.9	587	12	ADQ17664	Adq17664 Human sof	c 510	35	72.9	858	3	AAH70354	AAH70354 Human cer
438	35	72.9	593	4	ABA63443	AbA63443 Human foe	c 511	35	72.9	858	3	AAH70354	AAH70354 Human cer
439	35	72.9	593	4	ABA63443	AbA63443 Human foe	c 512	35	72.9	858	3	AAH70354	AAH70354 Human cer
440	35	72.9	593	4	ABA63443	AbA63443 Human foe	c 513	35	72.9	858	3	AAH70354	AAH70354 Human cer
441	35	72.9	593	4	ABA63443	AbA63443 Human foe	c 514	35	72.9	858	3	AAH70354	AAH70354 Human cer
442	35	72.9	593	4	ABA63443	AbA63443 Human foe	c 515	35	72.9	858	3	AAH70354	AAH70354 Human cer

516	35	72.9	1719	6	ABN84439	Abn84439 Drosophil	589	35	72.9	3864	6	ABO70993	Abq70993 Listeria
c 517	35	72.9	1748	13	ADX45519	Adx45519 Plant full	590	35	72.9	3970	4	ABO70782	Abt07782 Drosophil
518	35	72.9	1755	4	ABL06997	Abt06997 Drosophil	c 591	35	72.9	4058	4	ABU12932	Abt12932 Drosophil
519	35	72.9	1778	12	ADQ64801	Adq64801 Novel hum	592	35	72.9	4287	8	AAU597177	Aas97177 Human met
c 520	35	72.9	1788	13	ADX52070	Adx52070 Plant full	c 593	35	72.9	4290	8	AAU54359	Aal54359 Rat neuro
c 521	35	72.9	1794	8	ACA46702	Ac46702 Prokaryot	594	35	72.9	4359	2	AAT69969	Aat69969 FIVP epik
c 522	35	72.9	1803	6	ABN91336	Abn91336 Staphyloc	595	35	72.9	4359	14	ADY59204	Ady59204 Peline co
523	35	72.9	1814	4	ABL18419	Abt18419 Drosophil	596	35	72.9	4362	2	AAQ52440	Aq52440 Feline in
524	35	72.9	1845	2	AAQ07099	Aax07099 Staphyloc	597	35	72.9	4365	2	AAQ25536	Aq25536 S gene of
525	35	72.9	1852	11	ADM01462	Adm01462 Human cDN	598	35	72.9	4365	2	AAQ25534	Aq25534 S gene of
526	35	72.9	1872	4	ABL20173	Abt20173 Drosophil	599	35	72.9	4365	2	AAQ52443	Aq52443 Feline in
527	35	72.9	1895	5	ABA83001	Ab83001 Human tra	600	35	72.9	4365	2	AAQ52446	Aq52446 Feline in
c 528	35	72.9	1974	3	AAA70261	Aaa70261 Plasmodiu	c 601	35	72.9	4402	13	ACN42960	Acn42960 Human dia
c 529	35	72.9	1997	13	ADX59767	Adx59767 Plant full	602	35	72.9	4500	10	ACC43608	Acc43608 Nucleotid
530	35	72.9	2000	11	ACL35885	Ac135885 Rice stre	c 603	35	72.9	4585	10	ADE54472	Ad54472 Human gen
531	35	72.9	2029	10	ADB62720	Ad62720 Human cDN	c 604	35	72.9	4590	2	AAT94612	Aat94612 Rat penil
c 532	35	72.9	2067	4	AAF44622	Aaf44622 Novel pro	c 605	35	72.9	4768	4	ABL06996	Abt06996 Drosophil
c 533	35	72.9	2067	10	ACA56708	Ac56708 Human sig	606	35	72.9	4800	1	ABN81533	Abn81533 Sequence
c 534	35	72.9	2067	12	ADQ12803	Adq12803 Human GRK	c 607	35	72.9	4803	6	ABN83963	Abn83963 Human gen
c 535	35	72.9	2067	12	ADI29320	Adi29320 Human MAR	c 608	35	72.9	5063	12	ADK71887	Adk71887 Human kin
c 536	35	72.9	2067	12	ADI56504	Adi56504 Human pol	c 609	35	72.9	5066	14	ADX03746	Adx03746 Human mR
c 537	35	72.9	2067	14	ADV42680	Adv42680 Human psy	c 610	35	72.9	5108	2	AAQ53403	Aaq53403 Sequence
c 538	35	72.9	2115	12	ADN72952	Adn72952 Thale cre	611	35	72.9	5172	13	ADR31294	Adr31294 Aspergill
c 539	35	72.9	2186	5	ADI61327	Adi61327 Human ova	c 612	35	72.9	5246	14	ADX03718	Adx03718 Human mR
540	35	72.9	2189	6	ABQ69205	Abq69205 Listeria	c 613	35	72.9	5272	14	ADV97733	Adv97733 cDNA sequ
c 541	35	72.9	2233	2	AAQ02263	Aaq02263 cDNA sequ	c 614	35	72.9	5293	6	AAU44642	Aal44642 Orphan in
c 542	35	72.9	2233	10	ADG62817	Adg62817 Rat gene	c 615	35	72.9	5378	4	AAQ27810	Aas27810 DNA encod
543	35	72.9	2233	10	ADG62821	Adg62821 Rat gene	c 616	35	72.9	5378	10	ADB94613	Abd94613 Novel hum
544	35	72.9	2242	6	ADI28062	Adi28062 EWCAD ge	617	35	72.9	5471	4	ABU10094	Abt10094 Drosophil
545	35	72.9	2244	6	ABL66528	Abt66528 Lung canc	c 618	35	72.9	5512	4	ABU15242	Abt15242 Drosophil
546	35	72.9	2244	6	ABL65864	Abt65864 Lung canc	c 619	35	72.9	6090	2	AAT69977	Aat69977 FIVP modl
547	35	72.9	2244	13	ADU05836	Adu05836 Novel bro	c 620	35	72.9	6123	8	ADA71037	Ada71037 Rice gene
548	35	72.9	2244	14	ADY06145	Ady06145 Cyclin-de	621	35	72.9	6144	2	AAT69976	Aat69976 FIVP spik
549	35	72.9	2244	14	ADY17145	Ady17145 DNA encod	c 622	35	72.9	7075	2	AAV74574	Aav74574 Staphyloc
550	35	72.9	2246	2	AAQ25537	Aaq25537 Fragment	c 623	35	72.9	7792	4	ABL08222	Abt08222 Drosophil
551	35	72.9	2246	2	AAQ25541	Aaq25541 Consensus	c 624	35	72.9	8143	4	ABL09728	Abt09728 Drosophil
552	35	72.9	2246	2	AAQ25535	Aaq25535 A fragmen	625	35	72.9	8860	4	ABU06074	Abt06074 Drosophil
553	35	72.9	2246	2	AAQ52442	Aaq52442 Feline in	626	35	72.9	9278	2	AAQ76124	Aaq76124 Human MDC
554	35	72.9	2246	2	AAQ52441	Aaq52441 Feline in	c 627	35	72.9	10846	10	ADG70447	Adg70447 Continuation (5 of
555	35	72.9	2246	13	ACN37987	Acn37987 Tumour-as	c 628	35	72.9	10846	5	ABA21528	Ab21528 Human ner
556	35	72.9	2246	11	ADM02430	Adm02430 Human cDN	c 629	35	72.9	11860	4	ABL29524	Abt29524 Drosophil
c 557	35	72.9	2362	4	AAU58393	Aa158393 Human pol	c 630	35	72.9	12012	4	AAU29524	Aa29524 Human imm
c 558	35	72.9	2362	5	ADQ98603	Adq98603 DNA encod	c 631	35	72.9	12194	4	AAU28921	Aa28921 Human nov
c 559	35	72.9	2362	9	ADB48363	Abd48363 Novel hum	c 632	35	72.9	12194	10	ADB31762	Abd31762 Human complete
c 560	35	72.9	2373	13	ADV41500	Adv41500 Rat cardi	c 633	35	72.9	12249	3	AAU55840	Aa55840 Complete
c 561	35	72.9	2472	13	ADT17573	Adt17573 Plant cDN	c 634	35	72.9	13279	4	ABL20172	Abt20172 Drosophil
c 562	35	72.9	2501	6	ABZ78198	Abz78198 A. niger	635	35	72.9	13279	8	ABL20172	Abt20172 Drosophil
c 563	35	72.9	2526	12	ADQ67506	Adq67506 Novel hum	c 636	35	72.9	13789	8	ABZ73944	Abz73944 Secreted
564	35	72.9	2551	10	ADA52837	Ada52837 Human cod	c 637	35	72.9	13789	10	ADC20691	Adc20691 Human sec
c 565	35	72.9	2553	4	AAF84018	Aaf84018 Maize Rad	c 638	35	72.9	14197	8	ABX15607	Abx15607 FIVP plas
c 566	35	72.9	2568	8	ACA43326	Ac43326 Prokaryot	c 639	35	72.9	14482	4	ABL10086	Abt10086 Drosophil
567	35	72.9	2602	12	ADP64788	Adp64788 Drosophil	640	35	72.9	17258	4	ABL13486	Abt13486 Drosophil
568	35	72.9	2658	11	ADM03514	Adm03514 Human cDN	641	35	72.9	17844	3	AAA35236	Aa35236 Human ade
c 569	35	72.9	2667	4	ABL08223	Abt08223 Drosophil	642	35	72.9	17844	3	AAF21358	Aaf21358 Human low
570	35	72.9	2864	12	ADQ23035	Adq23035 Human sof	643	35	72.9	17844	10	ABZ97052	Abz97052 Human nuc
571	35	72.9	2864	12	ADQ23035	Adq23035 Human sof	644	35	72.9	17844	11	ABD20901	Abd20901 Human pul
c 572	35	72.9	2918	4	ABL25402	Abt25402 Drosophil	c 645	35	72.9	18331	3	AAU55857	Aa55857 Complete
c 573	35	72.9	2949	5	AAU67904	Aa67904 DNA encod	c 646	35	72.9	18331	10	ADE10266	Ad10266 S. lavend
c 574	35	72.9	2993	4	ABL09729	Abt09729 Drosophil	647	35	72.9	18331	3	AAU55857	Aa55857 Complete
c 575	35	72.9	2998	9	ABZ59024	Abz59024 Prunus se	648	35	72.9	18909	4	ABL07204	Abt07204 Drosophil
576	35	72.9	3199	13	ADR06670	Adr06670 Full leng	649	35	72.9	18981	4	ABL04961	Abt04961 Human rep
c 577	35	72.9	3237	14	ADZ49453	Adz49453 Insulin s	c 650	35	72.9	18981	4	ABL97854	Abt97854 Human tes
c 578	35	72.9	3268	4	ABL04134	Abt04134 Drosophil	c 651	35	72.9	28136	4	AAK69755	Aak69755 Human imm
c 579	35	72.9	3357	13	ADU66632	Adu66632 Human kin	c 652	35	72.9	28871	13	ADT05539	Adt05539 Haemophil
580	35	72.9	3362	4	ABU14286	Abt14286 Drosophil	c 653	35	72.9	42035	8	AAU59562	Aa59562 Propionib
581	35	72.9	3505	12	ADJ48200	Adj48200 Maize oil	654	35	72.9	42035	8	ACF64491	Acf64491 Propionib
c 582	35	72.9	3584	8	ABX63341	Abx63341 Human cDN	655	35	72.9	50833	12	ADQ97212	Adq97212 Mouse can
c 583	35	72.9	3628	14	ADX05742	Adx05742 Cyclin-de	656	35	72.9	53178	8	AAU59543	Aa59543 Propionib
584	35	72.9	3631	5	AAU05206	Aa05206 Drosophil	657	35	72.9	70242	11	ACN44710	Acn44710 Human gen
c 585	35	72.9	3670	4	AAU60179	Aa60179 Human pol	658	35	72.9	79122	3	AAU22294	Aa22294 BAC conta
586	35	72.9	3677	4	ABU15243	Abt15243 Drosophil	659	35	72.9	87394	13	ADT55151	Adt55151 Nucleotid
587	35	72.9	3741	4	ABL25388	Abt25388 Drosophil	660	35	72.9	87977	9	ADA02639	Ada02639 Mouse znf
588	35	72.9	3814	4	ABL18418	Abt18418 Drosophil	661	35	72.9	87977	10	ADB72377	Abd72377 Mouse znf

662	35	72.9	87977	10	AD895887	Ade95887 Mouse Znf	c 735	34	70.8	426	8	ABX53952	Abx53952 Bovine ES
663	35	72.9	110000	10	ADG70447_3	Continuation (4 of	736	34	70.8	430	2	ABX21100	Aax21100 Polynucle
664	35	72.9	110000	10	ABZ79565_3	Continuation (4 of	737	34	70.8	433	2	ABX48472	Ach48472 Bovine ES
665	35	72.9	110000	14	ADZ45062_00	Adz45062 Human cho	738	34	70.8	435	9	ACH49352	Ach49352 Human leu
666	35	72.9	110000	14	ABE42737_08	Continuation (9 of	739	34	70.8	437	6	ABL84102	Abl84102 Human ova
667	35	72.9	110000	14	ABE42737_08	Continuation (9 of	740	34	70.8	446	8	ABL47379	Abx47379 Bovine ES
668	35	72.9	110000	14	ABE42737_09	Continuation (10 o	741	34	70.8	450	13	ACF90615	Acf90615 Human SIR
669	35	72.9	111084	14	ADQ18808	Adq18808 Human sof	742	34	70.8	452	12	ADL13145	Adl13145 Human ste
670	35	72.9	115284	11	ACN44296	Acn44296 Mouse gen	743	34	70.8	460	6	ABL80027	Abl80027 Human ova
671	35	72.9	121124	12	ADQ97107	Adq97107 Mouse can	744	34	70.8	463	4	ABL54672	Abx54672 Human foe
672	35	72.9	128034	10	AD843582	Ade43582 Polymorph	745	34	70.8	463	4	AAI34329	Aai34329 Probe #30
673	35	72.9	128034	10	AD843581	Ade43581 Human IDE	746	34	70.8	463	4	ABA44221	Abx44221 Human bre
674	35	72.9	128034	12	ADH54059	Adh54059 Human IDE	747	34	70.8	463	4	ABA24454	Abx24454 Probe #29
675	35	72.9	128034	12	ADH54060	Adh54060 Human IDE	748	34	70.8	463	4	AAK28404	Aak28404 Human bon
676	35	72.9	178896	6	ABQ88146	Abq88146 Human oet	749	34	70.8	463	4	AAK02959	Aak02959 Human bra
677	35	72.9	188017	11	ACN45148	Acn45148 Mouse gen	750	34	70.8	463	6	ABS02913	Abx02913 Human gen
678	35	72.9	202100	10	ADH43315	Adh43315 Human IDE	751	34	70.8	465	9	ACH38837	Ach38837 Human foe
679	35	72.9	202100	12	ADH54357	Adh54357 Human IDE	752	34	70.8	471	6	ABL81275	Abl81275 Human ova
680	35	72.9	218802	14	ADW98820	Adw98820 Human her	753	34	70.8	476	9	ACH44848	Ach44848 Human foe
681	35	72.9	220756	12	ADG86300	Adg86300 Human SMR	754	34	70.8	483	4	AAK63951	Aak63951 Human imm
682	35	72.9	225883	13	ADV34981	Adv34981 Murine CD	755	34	70.8	484	6	ABL81865	Abl81865 Human ova
683	35	72.9	229354	6	ABQ74179	Abq74179 Human cyt	756	34	70.8	492	9	ACH13152	Ach13152 Human adu
684	35	72.9	233380	11	ACN44282	Acn44282 Human gen	757	34	70.8	492	10	ACD96379	Acd96379 Human col
685	35	72.9	234882	14	ADZ13715	Adz13715 Human can	758	34	70.8	496	13	ADU12572	Adu12572 Solid tum
686	35	72.9	246940	12	ADQ59422	Adq59422 Human can	759	34	70.8	516	12	ACH69011	Ach69011 Human gen
687	35	72.9	248436	11	ACN45190	Acn45190 Human gen	760	34	70.8	521	2	AAK21092	Aax21092 Polynucle
688	35	72.9	263853	14	ABE39171	Abx39171 L. pneumo	761	34	70.8	528	8	AXN91374	Axn91374 Murine ge
689	35	72.9	263853	14	ABE39171	Abx39171 L. pneumo	762	34	70.8	528	10	ABZ37970	Abz37970 N. gonorr
690	35	72.9	265118	5	AAH41227	Aah41227 Pyrococu	763	34	70.8	529	5	ADL39152	Adl39152 Human ova
691	35	72.9	294575	14	AEA61127	Aea61127 Human STK	764	34	70.8	533	2	AAK21099	Aax21099 Polynucle
692	35	72.9	308766	13	ADT05738	Adt05738 Haemophil	765	34	70.8	533	8	ACA04600	Aca04600 cDNA enco
693	35	72.9	313287	13	ABD33100	Abd33100 Human can	766	34	70.8	541	5	ADL45380	Adl45380 Human ova
694	35	72.9	347001	12	ADP33517	Adp33517 Human MAD	767	34	70.8	541	12	ADN12803	Adn12803 Human pro
695	35	72.9	349980	5	AAH68531	Aah68531 C Glutami	768	34	70.8	553	12	ACH79984	Ach79984 Human gen
696	34	70.8	23	2	AAT92409	Aat92409 Cosmid cl	769	34	70.8	557	9	AAK88420	Aak88420 Human dig
697	34	70.8	50	6	ABZ00030	Abz00030 Human leu	770	34	70.8	557	5	AAS39521	Aas39521 cDNA enco
698	34	70.8	50	6	ABZ00974	Abz00974 Human leu	771	34	70.8	557	9	ADB32247	Adb32247 Human nov
699	34	70.8	60	6	ABN36716	Abn36716 Human spl	772	34	70.8	557	13	ACN50076	Acn50076 Cotton pr
700	34	70.8	72	10	ACC72616	Acc72616 Human PCR	773	34	70.8	559	12	ACH74952	Ach74952 Human gen
701	34	70.8	135	13	ADS03603	Ads03603 Staphyloc	774	34	70.8	560	4	AAH11918	Aah11918 Human cDN
702	34	70.8	139	12	ACH93684	Ach93684 Human gen	775	34	70.8	564	12	ACH67239	Ach67239 Human pro
703	34	70.8	158	2	AAK12474	Aax12474 Human bia	776	34	70.8	570	12	ACH67239	Ach67239 Human gen
704	34	70.8	158	2	AAK12473	Aax12473 Human bia	777	34	70.8	571	13	ACN49151	Acn49151 Cotton pr
705	34	70.8	211	8	ABX55307	Abx55307 Bovine ES	778	34	70.8	582	4	ABA62899	Abx62899 Human foe
706	34	70.8	237	10	ACD93552	Acn93552 Human col	779	34	70.8	582	4	AAI42919	Aai42919 Probe #11
707	34	70.8	256	6	ABN75464	Abn75464 Human ORF	780	34	70.8	582	4	AAK37096	Aak37096 Human bon
708	34	70.8	276	2	AAV20322	Aav20322 Probe (22	781	34	70.8	582	4	AAK11303	Aak11303 Human bra
709	34	70.8	276	13	ADX10231	Adx10231 Plant ful	782	34	70.8	582	4	ABS36777	Abx36777 Human liv
710	34	70.8	277	12	ACH80944	Ach80944 Human gen	783	34	70.8	582	6	ABS11097	Abx11097 Human gen
711	34	70.8	283	3	AAK48394	Aac48394 Arabidops	784	34	70.8	583	3	AAK44618	Aac44618 Arabidops
712	34	70.8	285	12	ACH82711	Ach82711 Human gen	785	34	70.8	585	4	AAH10014	Aah10014 Human cDN
713	34	70.8	295	6	ABK77528	Abk77528 Bacillus	786	34	70.8	587	2	AAK21077	Aax21077 Polynucle
714	34	70.8	295	6	ABN94157	Abn94157 Gene #655	787	34	70.8	588	5	ABV55639	Abv55639 Human pro
715	34	70.8	324	6	ABQ98330	Abq98330 Human ORP	788	34	70.8	595	13	ACN59327	Acn59327 Cotton gy
716	34	70.8	331	4	AAK63600	Aak63600 Human inn	789	34	70.8	601	6	ABN63801	Abn63801 Human can
717	34	70.8	344	4	AAK85301	Aak85301 Human inn	790	34	70.8	602	2	AAK21076	Aax21076 Polynucle
718	34	70.8	345	4	AAK85303	Aak85303 Human inn	791	34	70.8	624	10	ACD94228	Acd94228 Human col
719	34	70.8	345	4	AAK85302	Aak85302 Human inn	792	34	70.8	637	6	ABI99727	Abi99727 Mouse isc
720	34	70.8	345	8	AAK54833	Abx54833 Bovine ES	793	34	70.8	659	6	ABS52781	Abx52781 Murine tu
721	34	70.8	346	2	ABL79903	Abi79903 Human ova	794	34	70.8	663	2	AAK21090	Aax21090 Polynucle
722	34	70.8	366	2	AAQ59742	Aac59742 Human bra	795	34	70.8	669	5	AAK64810	Aak64810 DNA enco
723	34	70.8	366	3	AAQ79819	Aac79819 Human sec	796	34	70.8	675	8	ACF74894	Acf74894 Staphyloc
724	34	70.8	401	4	AAK96261	Aak96261 Human neu	797	34	70.8	678	3	AAFI3815	Aafi3815 Aspergill
725	34	70.8	401	4	AAK97754	Aak97754 Human neu	798	34	70.8	678	8	ACA46362	Aca46362 Prokaryot
726	34	70.8	401	6	ABT01031	Abt01031 Human neu	799	34	70.8	678	13	ADU57856	Adu57856 Aspergill
727	34	70.8	401	6	ABT02524	Abt02524 Human neu	800	34	70.8	678	14	ADZ95859	Adz95859 Aspergill
728	34	70.8	402	6	ABL79493	Abi79493 Human ova	801	34	70.8	708	14	ADY56764	Ady56764 Human thr
729	34	70.8	407	9	ACH19747	Ach19747 Human adu	802	34	70.8	709	3	AAA48564	Aaa48564 Soybean p
730	34	70.8	408	5	AAK64809	Aak64809 DNA enco	803	34	70.8	720	6	ABT05593	Abt05593 Mycobacte
731	34	70.8	409	2	AAQ59429	Aaq59429 Human bra	804	34	70.8	720	8	ACA38701	Aca38701 Prokaryot
732	34	70.8	411	14	ADV75728	Adv75728 Human col	805	34	70.8	723	8	ACAA0352	Acaa0352 Prokaryot
733	34	70.8	416	12	ADB77187	Ade77187 Human cDN	806	34	70.8	725	8	ACC83307	Acc83307 Secreted
734	34	70.8	417	11	ABD07323	Abd07323 Pseudomon	807	34	70.8	725	9	ADA48795	Ada48795 Banana ge

C 808	34	70.8	725	11	ACL30866	Ac130866 Rice abio	881	34	70.8	1446	4	AA104154	Human rep
C 809	34	70.8	728	4	AA195329	Human neu	882	34	70.8	1446	5	AA540525	DNA encod
C 810	34	70.8	729	14	ADY56760	Human thr	883	34	70.8	1446	5	AA540528	DNA encod
C 811	34	70.8	729	14	ADY56766	Human thr	884	34	70.8	1446	5	AA540530	DNA encod
C 812	34	70.8	732	8	AD56070	Human sec	885	34	70.8	1446	11	ADJ09731	Human pro
C 813	34	70.8	732	14	ADY56758	Human thr	886	34	70.8	1446	11	ADJ09734	Human pro
C 814	34	70.8	732	14	ADY56774	Mouse thr	887	34	70.8	1446	11	ADJ09736	Human pro
C 815	34	70.8	732	14	ADY56762	Human thr	C 888	34	70.8	1458	2	AAx08528	NBP46 (ro
C 816	34	70.8	732	14	AD285226	Human Put	C 889	34	70.8	1458	2	AAx08528	NBP46 (ro
C 817	34	70.8	750	8	AC446732	Prokaryot	C 890	34	70.8	1458	6	ABK11099	DNA encod
C 818	34	70.8	753	8	ACC83303	Cotton an	C 891	34	70.8	1462	13	ABE65753	Rice geno
C 819	34	70.8	756	13	ACN54288	Adq34288 Cotton an	C 892	34	70.8	1476	13	AD548428	Bacterial
C 820	34	70.8	758	13	ADR64460	Adf64460 Cotton cD	C 893	34	70.8	1479	11	ACL26743	Rice abio
C 821	34	70.8	768	6	ABN91380	Abn91380 Staphyloc	C 894	34	70.8	1490	2	AAZ52932	Human pro
C 822	34	70.8	777	6	ABK75440	Abk75440 Bacillus	C 895	34	70.8	1490	5	ABV30210	Human pro
C 823	34	70.8	803	4	AA196426	AA196426 Human neu	C 896	34	70.8	1530	3	AAA96075	Potato ec
C 824	34	70.8	820	13	ADX65281	Adx65281 Plant ful	C 897	34	70.8	1530	10	ADJ57263	Potato ap
C 825	34	70.8	846	14	ABE65684	ABe65684 Rice geno	C 898	34	70.8	1533	13	AD545313	Bacterial
C 826	34	70.8	855	8	ACA27893	ACA27893 Prokaryot	C 899	34	70.8	1540	8	ABX72216	Human NOV
C 827	34	70.8	936	10	AD48798	Ad48798 Rhodococ	C 900	34	70.8	1545	11	ABD07287	Pseudomon
C 828	34	70.8	936	12	ADH18914	Adh18914 Human cel	C 901	34	70.8	1554	13	AD560096	Bacterial
C 829	34	70.8	955	13	ADQ38451	Adq38451 Human SNP	C 902	34	70.8	1554	13	ADX50536	Plant ful
C 830	34	70.8	975	3	ACS16128	ACS1628 Arabidops	C 903	34	70.8	1555	2	AAZ27325	Human sec
C 831	34	70.8	990	13	ADX53594	Adx53594 Plant ful	C 904	34	70.8	1555	8	ACC50474	Human sec
C 832	34	70.8	1005	5	ABD07357	ABd07357 Pseudomon	C 905	34	70.8	1555	8	ABZ71262	Human sec
C 833	34	70.8	1054	11	AA564468	AA564468 DNA encod	C 906	34	70.8	1555	9	ADA07204	Human cDN
C 834	34	70.8	1109	6	ABQ38147	ABq38147 Oligonuc1	C 907	34	70.8	1555	9	ADB91168	Human sec
C 835	34	70.8	1109	6	ABQ38146	ABq38146 Oligonuc1	C 908	34	70.8	1555	10	ADC73539	Human sec
C 836	34	70.8	1121	14	ADZ70553	Adz70553 Human cDN	C 909	34	70.8	1555	12	ADN40915	Novel hum
C 837	34	70.8	1125	6	ABS58379	ABs58379 Protein m	C 910	34	70.8	1577	10	AD31332	Human dia
C 838	34	70.8	1169	12	ADO20481	Ado20481 Human PRO	C 911	34	70.8	1581	4	ABA89512	Escherich
C 839	34	70.8	1177	13	ADX07068	Adx07068 Plant ful	C 912	34	70.8	1620	14	ACL68202	ACL68202 M. xanthu
C 840	34	70.8	1215	5	AAH67877	AAh67877 C Glutami	C 913	34	70.8	1621	4	ABL18465	ABL18465 Drosophil
C 841	34	70.8	1224	13	ADU00991	ADu00991 DNA encod	C 914	34	70.8	1622	4	ABL29923	ABL29923 Drosophil
C 842	34	70.8	1224	13	ADU15373	Adu15373 Novel hum	C 915	34	70.8	1624	13	ADX33014	Plant ful
C 843	34	70.8	1227	13	ACN38998	ACn38998 Tumour-as	C 916	34	70.8	1626	3	AAJ08691	PWAV-2 O
C 844	34	70.8	1227	13	ADXS2303	Adx52303 Plant ful	C 917	34	70.8	1646	8	ABZ74545	Secreted
C 845	34	70.8	1236	5	AA585318	AA585218 DNA encod	C 918	34	70.8	1646	10	ABZ68079	Human sec
C 846	34	70.8	1241	10	ADI60550	ADI60550 Secreted	C 919	34	70.8	1652	3	AACT79011	Human sec
C 847	34	70.8	1254	10	ADP53630	ADp53630 Haematopo	C 920	34	70.8	1652	8	ABZ73622	Secreted
C 848	34	70.8	1274	8	ABX10976	ABx10976 DNA seque	C 921	34	70.8	1652	10	ABZ67215	Human sec
C 849	34	70.8	1275	2	AA72871	AA72871 Gamma-smo	C 922	34	70.8	1652	12	ADJ57902	Human PGP
C 850	34	70.8	1275	11	AD131482	Ad131482 Human cDN	C 923	34	70.8	1677	3	AAZ3066	Arabidops
C 851	34	70.8	1275	13	AD583549	AD583549 Human lym	C 924	34	70.8	1682	4	ABL11859	ABL11859 Drosophil
C 852	34	70.8	1280	12	ADY19756	ADy19756 Human PRO	C 925	34	70.8	1683	6	ABK69118	DNA encod
C 853	34	70.8	1280	14	ADY16388	ADy16388 DNA encod	C 926	34	70.8	1717	10	ADJ22615	Human liv
C 854	34	70.8	1288	12	ADP45416	Adf45416 Human vas	C 927	34	70.8	1718	4	AAH33518	Human col
C 855	34	70.8	1288	12	ADN03844	Adn03844 Antipsori	C 928	34	70.8	1743	9	ACD26508	Human pro
C 856	34	70.8	1288	13	ADR24777	Adr24777 Breast ca	C 929	34	70.8	1743	10	ACH00821	Human pro
C 857	34	70.8	1288	14	ADY54924	ADy54924 Chronic v	C 930	34	70.8	1757	4	AAI58424	Human pol
C 858	34	70.8	1299	13	ADS58544	Ad58544 Bacterial	C 931	34	70.8	1757	5	ADQ98634	DNA encod
C 859	34	70.8	1302	8	ACA23121	ACA23121 Prokaryot	C 932	34	70.8	1757	9	AD848394	Novel hum
C 860	34	70.8	1302	13	ADU98768	Adu98768 Borrelia	C 933	34	70.8	1763	2	AAZ21163	Artificia
C 861	34	70.8	1304	13	ADX29202	Adx29202 Plant ful	C 934	34	70.8	1780	14	ADZ60324	Murine AC
C 862	34	70.8	1312	6	ABN98247	ABn98247 Arabidops	C 935	34	70.8	1787	2	AAZ21168	Artificia
C 863	34	70.8	1313	13	ADT16423	Adt16423 Plant cDN	C 936	34	70.8	1791	12	ADM87458	Human EST
C 864	34	70.8	1317	4	AAZ60275	AAz60275 Human ERA	C 937	34	70.8	1792	6	ABK69092	DNA encod
C 865	34	70.8	1333	14	ABE27061	ABe27061 Pinus rad	C 938	34	70.8	1797	13	ADU07848	DNA seque
C 866	34	70.8	1338	4	AA71261	AA71261 Corynebac	C 939	34	70.8	1802	13	ADT16754	Plant cDN
C 867	34	70.8	1354	4	ABL17123	ABl17123 Drosophil	C 940	34	70.8	1805	2	AAZ21165	Artificia
C 868	34	70.8	1356	8	ACC83306	ACC83306 Secreted	C 941	34	70.8	1809	14	ACL71299	M. xanthu
C 869	34	70.8	1356	13	ADSI0211	ADs10211 Human the	C 942	34	70.8	1840	4	AAH16366	Human cDN
C 870	34	70.8	1362	10	ADI60478	Adi60478 Human mdd	C 943	34	70.8	1841	4	AAI61240	Human pol
C 871	34	70.8	1372	8	ABX34776	ABx34776 Human mdd	C 944	34	70.8	1841	10	ADI60276	Secreted
C 872	34	70.8	1375	5	AA585219	AA585219 DNA encod	C 945	34	70.8	1845	2	AAZ89289	Human reg
C 873	34	70.8	1376	10	ADBA7455	ADb7455 Human cDN	C 946	34	70.8	1846	4	AAI60210	Human pol
C 874	34	70.8	1381	3	AAA48572	AAA48572 Soybean p	C 947	34	70.8	1861	4	AAH17210	Human cDN
C 875	34	70.8	1386	5	AA68721	AA68721 DNA encod	C 948	34	70.8	1861	10	ADD18975	Human dis
C 876	34	70.8	1404	4	AAH98660	AAh98660 Human EST	C 949	34	70.8	1872	13	ADM47560	Leptospir
C 877	34	70.8	1419	13	ADR26906	Adr26906 Rat granu	C 950	34	70.8	1873	3	AAZ99901	Human sec
C 878	34	70.8	1441	10	ADBA7456	ADb7456 Human cDN	C 951	34	70.8	1888	14	ADZ07763	HIV CON-C
C 879	34	70.8	1446	4	AA104156	AA104156 Human rep	C 952	34	70.8	1891	14	ADZ07757	HIV CON-S
C 880	34	70.8	1446	4	AA104151	AA104151 Human rep	C 953	34	70.8	1899	10	ABZ38070	N. gonorr

PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.

XX Example 9; Page 302; 959pp; English.

XX The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC cDNA, sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleotide sequence of an open reading frame; for preparing a contig,
CC nucleic acid molecule from a genome of an organism; and for sequencing
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a fetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library

SQ Sequence 299 BP; 75 A; 84 C; 78 G; 62 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.03 Length: 299
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-9 (1-9) x ACD93536 (1-299)

QY 1 GlyLeuProHisIleArgValPheLeu 9
Db 141 GGTCTACCCACATTAGGCTTTTCCTG 167

RESULT 3
ID AAS87174

XX AAS87174 standard; cDNA; 453 BP.

AC AAS87174;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #22978.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG22987.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 1; SEQ ID NO 22978; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 453 BP; 108 A; 111 C; 113 G; 121 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.3 Length: 453
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-9 (1-9) x AAS87174 (1-453)

QY 1 GlyLeuProHisIleArgValPheLeu 9
Db 34 GGTCTACCCACATTAGGCTTTTCCTG 60

RESULT 4

ID ADU11677

XX ADU11677 standard; DNA; 475 BP.

AC ADU11677;

XX 27-JAN-2005 (first entry)

DE Solid tumour prognosis gene segid 2116.

XX cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.

XX Unidentified.

XX

PN WO2004097052-A2.
XX 11-NOV-2004.
XX 29-APR-2004; 2004WO-US013587.
XX 29-APR-2003; 2003US-0466067P.
PR 23-JAN-2004; 2004US-0538246P.
XX (AMHP) WYETH.
PA (STRA/) STRAHS A.
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dorner AJ;
DR WPI; 2004-804779/79.
XX
XX A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
XX Disclosure; Page; ilpp; English.
XX
XX The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumor, and each of the gene is differentially
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumor, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumor, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumor, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumor correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumor prognosis genes, where
CC each of the solid tumor prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumor, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumors. This sequence represents a solid tumor prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores: 3.48 Length: 475
Pred. No.: 48.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 13 Gaps: 0
DB:

US-10-774-176-9 (1-9) x ADU11677 (1-475)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
|||||
Db 126 GGTCTACCCACATTAGGGTTTCCTG 152

RESULT 5
ABT07721
ID ABT07721 standard; DNA; 927 BP.
XX
AC ABT07721;
XX
DT 14-NOV-2002 (first entry)
XX
DE Breast cancer-associated gene sequence 29.
XX
KW Gene; ds; breast cancer; breast cancer-associated gene sequence;
KW drug development; pharmacogenetics; biosensor development.
XX
OS Unidentified.
XX
PN WO200259377-A2.
XX
PD 01-AUG-2002.
XX
PF 24-JAN-2002; 2002WO-US002242.
XX
PR 24-JAN-2001; 2001US-0263965P.
PR 02-FEB-2001; 2001US-0265928P.
PR 09-APR-2001; 2001US-00829472.
PR 09-APR-2001; 2001US-0282698P.
PR 04-MAY-2001; 2001US-0288590P.
PR 29-MAY-2001; 2001US-0294443P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Gish KC, Afar D;
XX
DR WPI; 2002-583738/62.
DR N-PSDB; AB055564.
XX

PT Detecting a breast cancer-associated transcript in a patient's cell,
PT useful for diagnosing breast cancer, comprises contacting a biological
PT sample with a polynucleotide that selectively hybridizes with breast
PT cancer nucleic acids.
XX
PS Claim 9; Page 372; 414pp; English.
XX
CC The invention comprises a method of detecting a breast cancer-associated
CC transcript in a cell from a patient. The method of the invention involves
CC contacting a biological sample from the patient with a nucleotide that
CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
CC in the specification. The method of the invention is useful in the
CC diagnosis or prognosis of breast cancer, and for detecting genes that are
CC up or down-regulated in breast cancer cells. Genes identified by the
CC method of the invention can be used in diagnostic purposes and also as
CC targets for screening for therapeutic compounds that modulate breast
CC cancer (e.g. hormones or antibodies). Identification of genes that are
CC over or under expressed in breast cancer can additionally provide high-
CC resolution, high-sensitivity datasets which can be used in the areas of
CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
CC structure and biosensor development. DNA sequences AB076931, AB077761
CC represent the 69 breast cancer-associated gene sequences of the invention
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores: 7.59 Length: 927
Pred. No.: 48.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 6 Gaps: 0
DB:

US-10-774-176-9 (1-9) x ABT07721 (1-927)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
|||||
Db 505 GGTCTACCCACATTAGGGTTTCCTG 531

RESULT 6
ABX76333
ID ABX76333 standard; DNA; 927 BP.
XX AC ABX76333;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polynucleotide #197.
XX KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US012476.
XX PR 18-APR-2001; 2001US-0284770P.
XX PR 10-MAY-2001; 2001US-0290492P.
XX PR 09-NOV-2001; 2001US-0339245P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 29-NOV-2001; 2001US-0334370P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
XX DR P-PSDB; ABUS6604.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
XX PT for treating lung cancer, by contacting a biological sample from the
XX PT patient with a polynucleotide that exhibits increased or decreased
XX PT expression in lung cancer.
XX PS Claim 22; Page 336; 453pp; English.
XX CC The invention relates to a method for detecting a lung cancer-associated
XX CC transcript in a cell from a patient, comprising contacting a biological
XX CC sample from the patient with a polynucleotide that selectively hybridizes
XX CC to a sequence that is at least 80 % identical to a gene that exhibits
XX CC increased or decreased expression in lung cancer samples. Lung cancer-
XX CC associated polynucleotides and polypeptides are used for identifying a
XX CC compound that modulates a lung cancer-associated polypeptide, for
XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung
XX CC cancer in a patient and for treating a mammal having lung cancer by
XX CC administering a modulatory compound identified. The methods are useful
XX CC for treating lung cancer, such as small cell lung cancer, non-small cell
XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX CC for diagnostic purposes and as targets for screening for therapeutic
XX CC compounds that modulate lung cancer, such as antibodies. Sequences
XX CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
XX CC invention
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 7.59 Length: 927
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-10-774-176-9 (1-9) x ABX76333 (1-927)
QY 1 GlyLeuProHisIleArgValPheLeu 9
DB 505 GGTCTACCCACATTAGGTTTCCTG 531
RESULT 7
ADB80503
ID ADB80503 standard; DNA; 927 BP.
XX AC ADB80503;
XX DT 04-DEC-2003 (first entry)
XX DE Ovarian cancer-associated transcript #34.
XX KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection; ds; gene.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT CDS 1..927
XX FT /*tag= a
XX PN WO2002102235-A2.
XX PD 27-DEC-2002.
XX PF 18-JUN-2002; 2002WO-US019297.
XX PR 18-JUN-2001; 2001US-0299234P.
XX PR 27-AUG-2001; 2001US-0315287P.
XX PR 05-SEP-2001; 2001US-0317544P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Mack DH, Gish KC;
XX DR WPI; 2003-167431/16.
XX DR P-PSDB; ADB80504.
XX CC Detecting an ovarian cancer-associated transcript in a cell from a
XX CC patient, comprises contacting a biological sample from the patient with a
XX CC polynucleotide that hybridizes to an ovarian cancer gene.
XX PS Claim 10; Page 297; 332pp; English.
XX CC The invention relates to a method of detecting an ovarian cancer-
XX CC associated transcript in a cell from a patient, by contacting a
XX CC biological sample from the patient with a polynucleotide that selectively
XX CC hybridizes to a sequence at least 80% identical to any of one of 80
XX CC nucleic acid sequences given in the specification. The method is useful
XX CC in diagnosing ovarian cancer and in identifying and using agents and/or
XX CC targets that inhibit ovarian cancer. The nucleic acid molecule,
XX CC polypeptide and the antibody may also be used in detecting ovarian
XX CC cancers, monitoring and early detection of relapse following treatment,
XX CC monitoring response to therapy, selecting patients for post-operative
XX CC chemotherapy or radiation therapy, in selecting mode of therapy,
XX CC determining tumour prognosis, early detection of pre-cancerous lesions,
XX CC and as vaccines. This sequence corresponds to one of the nucleic acids
XX CC used for the detection method of the invention.
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 7.59 Length: 927

Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-9 (1-9) x ADB80503 (1-927)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 505 GGTCTACCCACATTAGGGTTTCTG 531

RESULT 8

ID ADN38723

ADN38723 standard; cDNA; 927 BP.

XX AC ADN38723;

XX DT 17-JUN-2004 (first entry)

XX DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX KW Human, differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnary; gene therapy; vaccine; gene; ss.
 XX OS Homo sapiens.

XX FN WO2003042661-A2.

XX PD 22-MAY-2003.

XX PF 13-NOV-2002; 2002WO-US036810.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 21-NOV-2001; 2001US-0332464P.

XX PR 29-NOV-2001; 2001US-0334393P.

XX PR 03-DEC-2001; 2001US-0335394P.

XX PR 14-DEC-2001; 2001US-0340376P.

XX PR 08-JAN-2002; 2002US-0347211P.

XX PR 10-JAN-2002; 2002US-0347349P.

XX PR 08-FEB-2002; 2002US-0355250P.

XX PR 13-FEB-2002; 2002US-0356714P.

XX PR 20-FEB-2002; 2002US-0359077P.

XX PR 29-MAR-2002; 2002US-0358809P.

XX PR 04-APR-2002; 2002US-0370110P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PR 05-JUN-2002; 2002US-0386614P.

XX PR 16-JUL-2002; 2002US-0396839P.

XX PR 22-JUL-2002; 2002US-039775P.

XX PR 22-JUL-2002; 2002US-0397845P.

XX PR 09-SEP-2002; 2002US-0409450P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.

XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.59 Length: 927
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

US-10-774-176-9 (1-9) x ADN38723 (1-927)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 505 GGTCTACCCACATTAGGGTTTCTG 531

RESULT 9

AAD56198

ID AAD56198 standard; DNA; 973 BP.

XX AC AAD56198;

XX DT 07-AUG-2003 (first entry)

XX DE Human LRRCAPS related DNA #5.

XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.

XX OS Homo sapiens.

XX FN WO2003035831-A2.

XX PD 01-MAY-2003.

XX PF 21-OCT-2002; 2002WO-US033540.

XX PR 22-OCT-2001; 2001US-0338733P.

XX PR 15-FEB-2002; 2002US-0357600P.

XX PR 01-MAR-2002; 2002US-0361196P.

XX PA (EXEL-) EXELIXIS INC.

XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;

XX PI Francis-Lang H, Friedman L;

XX DR WPI; 2003-421410/39.

XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 XX comprises contacting an assay system comprising a purified leucine rich
 XX repeat, capricious related polypeptide or nucleic acid with a test agent.

XX PS Example 5; Page 74-75; 99pp; English.

XX CC The invention relates to a method of identifying a candidate p53 pathway

XX CC modulating agent. The method involves contacting an assay system

XX CC comprising a purified Leucine rich repeat, capricious related (LRRCAPS)

WPI; 2003-468649/44.

P-PSDB; ADN38724.

Determining the presence or absence of a pathological cell in a patient,
 useful for diagnosing, prognosing or treating cancer, comprises detecting
 a nucleic acid in a biological sample.

Claim 8; SEQ ID NO 41; 1385pp; English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 whose expression is upregulated or downregulated in specific cancers or

CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX
 SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.03 Length: 973
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x AAD56198 (1-973)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 520 GGTCTACCCACATTAGGGTTTCTCG 546

RESULT 10

ABV99349
 ID ABV99349 standard; DNA; 1156 BP.

AC ABV99349;

XX 27-JAN-2003 (first entry)

XX Human NOV8a coding sequence.

XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
 KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

XX Homo sapiens.

XX WO200272771-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US007288.

XX 08-MAR-2001; 2001US-0274101P.

XX 08-MAR-2001; 2001US-0274194P.

XX 08-MAR-2001; 2001US-0274281P.

XX 09-MAR-2001; 2001US-0274322P.

XX 12-MAR-2001; 2001US-0274849P.

XX 13-MAR-2001; 2001US-0275235P.

XX 13-MAR-2001; 2001US-0275578P.

XX 13-MAR-2001; 2001US-0275601P.

XX 14-MAR-2001; 2001US-0276000P.

XX 16-MAR-2001; 2001US-0276766P.

XX 19-MAR-2001; 2001US-0276994P.

XX 20-MAR-2001; 2001US-0277239P.

XX 20-MAR-2001; 2001US-0277321P.

PR 27-MAR-2001; 2001US-0278999P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0279959P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 02-APR-2001; 2001US-0280900P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 02-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 03-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 30-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332722P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 03-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.

(CURA-) CURAGEN CORP.

Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 Zhong M;

WPI; 2002-732824/79.

P-PSDB; ABP70071.

New NOVX polypeptides and polynucleotides, useful for preventing,
 diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 disorders, and asthma.

Claim 16; Page 114-115; 619pp; English.

The present invention relates to new isolated proteins (NOVX) and their
 coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 any number from 1 to 48. The NOVX proteins and coding sequences are
 useful in the manufacture of a medicament for treating a syndrome
 associated with a human disease, preferably a NOVX-associated disorder.
 The NOVX coding sequences and proteins are useful for treating, diabetes,
 preventing or diagnosing diseases such as metabolic disorders, diabetes,
 obesity, infectious disease, anorexia, cancer-associated cachexia,
 cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 disease, immune disorders, haematopoietic disorders, cardiovascular

CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods

XX Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 9.82 Length: 1156
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-9 (1-9) x ABV99349 (1-1156)

QY 1 GlyLeuProHisIleArgValPheLeu 9
DB 736 GGTCTACCCACATTAGGGTTTCCTG 762

RESULT 11
AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX AC AAA27058;
XX 22-AUG-2000 (first entry)
XX Human 5T4 tumour-associated antigen gene.
XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX ds.
XX Homo sapiens.
XX WO200029428-A2.
XX 25-MAY-2000.
XX 18-NOV-1999; 99WO-GB003859.
XX 18-NOV-1998; 98GB-00025303.
XX 27-JAN-1999; 99GB-00001739.
XX 30-JUL-1999; 99GB-00017995.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX Carroll MW, Myers KA;
XX WPI; 2000-387735/33.
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX response useful in vaccinating against and in treating tumors.
XX Example 2; Page 78; 79pp; English.
XX The present sequence encodes the human 5T4 tumour-associated antigen
XX (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX carcinomas but has a highly restricted expression pattern in normal adult
XX tissues. It appears to be strongly correlated to metastasis in colorectal
XX and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX induced were inoculated with a virus expression vector containing the
XX present sequence. The 5T4 antigen was shown to be effective at eliciting
XX an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX the antigen and the antigen itself can be used to elicit an immune

CC response, preferably CTL or an antibody response in a subject

XX Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 10.9 Length: 1263
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-9 (1-9) x AAA27058 (1-1263)

QY 1 GlyLeuProHisIleArgValPheLeu 9
DB 847 GGTCTACCCACATTAGGGTTTCCTG 873

RESULT 12
AAD56199
ID AAD56199 standard; DNA; 1331 BP.
XX AC AAD56199;
XX 07-AUG-2003 (first entry)
XX Human LRRCAPS related DNA #6.
XX Human; p53 pathway; Leucine rich repeat capricious related protein;
XX LRRCAPS; cancer; gene therapy; ds.
XX Homo sapiens.
XX WO2003035831-A2.
XX 01-MAY-2003.
XX 21-OCT-2002; 2002WO-US033540.
XX 22-OCT-2001; 2001US-0338733P.
XX 15-FEB-2002; 2002US-0357600P.
XX 01-MAR-2002; 2002US-0361196P.
XX (EXEL-) EXELIXIS INC.
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
XX Francis-Lang H, Friedman L;
XX WPI; 2003-421410/39.
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
XX comprises contacting an assay system comprising a purified leucine rich
XX repeat, capricious related polypeptide or nucleic acid with a test agent.
XX Disclosure; Page 75-76; 99pp; English.
XX The invention relates to a method of identifying a candidate p53 pathway
XX modulating agent. The method involves contacting an assay system
XX comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX polypeptide or nucleic acid or its fragment with a test agent and
XX detecting a test agent-biased activity, where a difference between the
XX test agent-biased activity and the reference activity identifies the test
XX agent as a candidate p53 pathway modulating agent. The method is useful
XX for identifying a candidate p53 pathway-modulating agent for preparing a
XX composition for diagnosing or treating cancer. The invention is useful in
XX gene therapy. The present sequence is human LRRCAPS related DNA

XX Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 11.6 Length: 1331
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x AAD56199 (1-1331)

QY 1 GlyLeuProHisIleArgValPheLeu 9
DB 877 GGTCTACCCACATAGGGTTTCCTG 903

RESULT 13

ADJ56299
ID ADJ56299 standard; cDNA; 2020 BP.

XX AC ADJ56299;

XX DT 06-MAY-2004 (first entry)

XX DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.

XX KW human; differential expression; transactivator; proto-oncogene;

XX KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;

XX KW MYCN activated cell.

XX OS Homo sapiens.

XX PN US2003119009-A1.

XX PD 26-JUN-2003.

XX PF 25-FEB-2002; 2002US-00084817.

XX PR 23-FEB-2001; 2001US-0270784P.

XX PA (STUA/) STUART S G.

XX PA (NUCH/) NUCHTERN J G.

XX PA (PLON/) PLON S E.

XX PA (SHOH/) SHOHET J M.

XX PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;

XX DR WPI; 2003-635698/60.

XX PT New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.

PS Claim 1; SEQ ID NO 105; 27pp; English.

XX CC This invention relates to novel isolated cDNAs that are differentially
XX CC expressed in MYCN activated cells. Specifically, it refers to
XX CC polynucleotide sequences that exhibit differential expression patterns in
XX CC cells activated by the transactivator MYCN, where MYCN is a proto-
XX CC oncogene that is amplified in neuroblastoma cells and is common in small
XX CC cell lung cancers. The present invention describes these cDNA molecules
XX CC as useful for in hybridisation assays to detect expression of nucleic
XX CC acids (or complementary nucleic acids) in a present in a given sample, as
XX CC well as for screening assays by identifying molecules or compounds that
XX CC specifically bind the cDNA as a ligand and modulate function or activity.
XX CC Accordingly, these compositions exhibit cytostatic activity and can also
XX CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX CC that is differentially expressed in MYCN activated cells, given in an
XX CC exemplification of the invention. NOTE: This sequence does not appear in
XX CC the printed specification but has been obtained in electronic format from
XX CC the US Patent Office at
XX CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.

SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 18.8 Length: 2020
Score: 48.00 Matches: 9

Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-9 (1-9) x ADJ56299 (1-2020)

QY 1 GlyLeuProHisIleArgValPheLeu 9
DB 917 GGTCTACCCACATAGGGTTTCCTG 943

RESULT 14

ACCS1052

ID ACCS1052 standard; cDNA; 2053 BP.

XX AC ACCS1052;

XX DT 12-JUN-2003 (first entry)

XX DE Human bladder cancer associated cDNA sequence SEQ ID NO:192.

XX KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX PN WO2003003906-A2.

XX PD 16-JAN-2003.

XX PF 03-JUL-2002; 2002WO-US021338.

XX PR 03-JUL-2001; 2001US-0302814P.

XX PR 03-AUG-2001; 2001US-0310099P.

XX PR 08-NOV-2001; 2001US-0343705P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Mack DH, Aziz N;

XX DR WPI; 2003-201532/19.

XX DR P-PSDB; ABR48236.

XX PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.

PS Claim 6; Page 296; 307pp; English.

XX CC The present invention describes a method for detecting a bladder cancer-
XX CC associated transcript in a cell from a patient. The method comprises
XX CC contacting a biological sample from the patient with a polynucleotide
XX CC that selectively hybridises to a sequence that is 80 % identical to a
XX CC table of sequences (see ACCS0951 to ACCS1059). ACCS0951 to ACCS1059
XX CC encode the human bladder cancer-associated proteins given in ABR48146 to
XX CC ABR48242). Bladder cancer-associated sequences from the present invention
XX CC have cytostatic activities, and can be used in antisense gene therapy and
XX CC in vaccine production. The method can be used for detecting a bladder
XX CC cancer-associated transcript in a cell from a patient. The method is
XX CC useful in diagnosing or treating bladder cancer and in screening for
XX CC compounds that modulate bladder cancer, such as hormones or antibodies.
XX CC The nucleic acid molecules from the present invention may be used in
XX CC various screening and diagnostic methods, and for gene therapy, vaccine
XX CC and/or antisense/inhibition applications

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 19.2 Length: 2053
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0


```
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x ACC51052 (1-2053)
Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 931 GGCTACCCACATTTAGGGTTTCCTG 957

RESULT 15
ABX76332
ID ABX76332 standard; DNA; 2053 BP.
XX
AC ABX76332;
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polynucleotide #196.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
FN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
DR P-P8DB; ABU56603.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
XX
XX Claim 22; Page 335; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention
```

```
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 19.2 Length: 2053
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x ABX76332 (1-2053)
Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 931 GGCTACCCACATTTAGGGTTTCCTG 957

RESULT 16
AAD56197
ID AAD56197 standard; DNA; 2053 BP.
XX
AC AAD56197;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human LRRCAPS DNA #11.
XX
KW Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
FN WO2003035831-A2.
XX
PD 01-MAY-2003.
XX
PF 21-OCT-2002; 2002WO-US033540.
XX
PR 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357600P.
PR 01-MAR-2002; 2002US-0361196P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
XX WPI; 2003-421410/39.
XX
PT Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
XX Example 5; Page 73-74; 99pp; English.
XX
CC The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS DNA
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 19.2 Length: 2053
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
```

Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x AAD56197 (1-2053)

QY 1 GlyLeuProHisIleArgValPheLeu 9
 |||||
 DB 931 GGTCTACCCACATTAGG3GTTTCCTG 957

RESULT 17

AAD56200
 ID AAD56200 standard; DNA; 2053 BP.
 XX AC AAD56200;
 XX DT 07-AUG-2003 (first entry)
 XX DE Human LRRCAPS DNA #12.
 XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; Gene therapy; ds.
 XX OS Homo sapiens.
 XX PN WO2003035831-A2.
 XX PD 01-MAY-2003.
 XX PF 21-OCT-2002; 2002WO-US033540.
 XX PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX PA (EXEL-) EXELIXIS INC.
 XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX WPI; 2003-421410/39.
 XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX PS Disclosure; Page 76-77; 99pp; English.
 XX CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 19.2 Length: 2053
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x AAD56200 (1-2053)

QY 1 GlyLeuProHisIleArgValPheLeu 9
 |||||
 DB 931 GGTCTACCCACATTAGG3GTTTCCTG 957

RESULT 18

ADN38721
 ID ADN38721 standard; cDNA; 2053 BP.
 XX AC ADN38721;
 XX DT 17-JUN-2004 (first entry)
 XX DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.
 XX KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ds.
 XX OS Homo sapiens.
 XX PN WO2003042661-A2.
 XX PD 22-MAY-2003.
 XX PF 13-NOV-2002; 2002WO-US036810.
 XX PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX WPI; 2003-468649/44.
 XX DR P-PSDB; ADN38722.
 XX PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX PS Claim 8; SEQ ID NO 39; 1385pp; English.
 XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,

CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 19.2 Length: 2053
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-9 (1-9) x ADN38721 (1-2053)

QY 1 GlyLeuProHisIleArgValPheLeu 9
|||||
DB 931 GGTCTACCCACATTAGGGTTTCCTG 957

RESULT 19

ADL06473
ID ADL06473 standard; cDNA; 2053 BP.

AC ADL06473;

XX 20-MAY-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #53.

XX Human; tumour-associated antigenic target; TAT; cell death; tumour;
KW cancer; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004016225-A2.

XX 26-FEB-2004.

XX 19-AUG-2003; 2003WO-US025892.

XX 19-AUG-2002; 2002US-0404809P.

XX 21-AUG-2002; 2002US-0405645P.

XX 23-SEP-2002; 2002US-0413192P.

XX 15-OCT-2002; 2002US-0419008P.

XX 15-NOV-2002; 2002US-0426847P.

XX 02-JUL-2003; 2003US-0484959P.

XX (GETH) GENENTECH INC.

XX Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;

XX Spencer SD, Wu TD, Zhang Z;

XX WPI; 2004-257144/24.

XX P-PSDB; ADL06552.

XX New antibody that binds to a tumor-associated antigenic target (TAT)
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating cancer.

XX Claim 1; SEQ ID NO 53; 319pp; English.

XX The present invention relates to the isolation of human tumour-associated
XX antigenic target (TAT) polynucleotide and polypeptide sequences. Also
XX disclosed is an antibody that binds to a TAT polypeptide. The antibody is
XX a monoclonal antibody, an antibody fragment, a chimeric antibody or a
XX humanised antibody. It is conjugated to a growth inhibitory agent. It is
XX produced in bacteria or in CHO cells and induces death of a cell to which
XX it binds. The antibody is useful for preparing a composition for
XX diagnosing or treating tumours and cancer. The present sequence
XX represents a human TAT cDNA sequence of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 19.2 Length: 2053
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-9 (1-9) x ADL06473 (1-2053)

QY 1 GlyLeuProHisIleArgValPheLeu 9
|||||
DB 931 GGTCTACCCACATTAGGGTTTCCTG 957

RESULT 20

ADN03961

ID ADN03961 standard; cDNA; 2053 BP.

XX AC ADN03961;

XX 01-JUL-2004 (first entry)

XX Antipsoriatic cDNA sequence #180.

XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

XX WO2004028479-A2.

XX 08-APR-2004.

XX 25-SEP-2003; 2003WO-US030907.

XX 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;

XX WPI; 2004-305105/28.

XX P-PSDB; ADN03962.

XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.

XX Claim 1; SEQ ID NO 355; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 19.2 Length: 2053
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-9 (1-9) x ADN03961 (1-2053)

QY 1 GlyLeuProHisIleArgValPheLeu 9
|||||
DB 931 GGTCTACCCACATTAGGGTTTCCTG 957

RESULT 21

ADR25444

ID ADR25444 standard; DNA; 2053 BP.

XX AC ADR25444;

XX DT 21-OCT-2004 (first entry)

XX DE Breast cancer prognosis marker #1305.

XX ds; breast cancer; prognosis; gene expression; diagnosis.

XX OS Homo sapiens.

XX FN WO2004065545-A2.

XX PD 05-AUG-2004.

XX PF 15-JAN-2004; 2004WO-US001100.

XX PR 15-JAN-2003; 2003US-00342887.

XX PA (ROSE-) ROSETTA INPHARMATICS LLC.

XX PA (NECA-) NETHERLANDS CANCER INST.

XX PI Van't Veer LJ, He Y;

XX DR WPI; 2004-593473/57.

XX PT Classifying a breast cancer patient according to prognosis comprises

XX PT determining the similarity between the level of expression of each of

XX PT five genes in a cell sample taken from patient, to control levels.

XX PS Disclosure; SEQ ID NO 1305; 226pp; English.

XX CC The invention relates to a method of classifying a breast cancer patient

XX CC according to prognosis by determining the similarity between the level of

XX CC expression of each of five genes for which markers are listed in the

XX CC specification, in a cell sample taken from the breast cancer patient, to

XX CC control levels of expression for each respective five genes to obtain a

XX CC patient similarity value. The methods are useful for classifying a breast

XX CC cancer patient according to prognosis. Kits and computer program products

XX CC are useful for data analysis using the diagnostic, prognostic and

XX CC statistical methods of the invention. This sequence corresponds to a

XX CC marker used in the method of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

KW

KW

KW

KW

KW

KW

KW

XX

XX

OS

XX

XX

FN

XX

XX

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Tumour-associated antigenic target; TAT; human; overexpression; cancer; tumour; diagnosis; cell proliferative disorder; breast cancer; colorectal cancer; lung cancer; ovarian cancer; liver cancer; central nervous system cancer; bladder cancer; pancreatic cancer; cervical cancer; melanoma; leukaemia; hybridisation probe; chromosome identification; chromosome mapping; gene mapping; gene therapy; cytostatic; gene; ss.

Homo sapiens.

WO2004030615-A2.

15-APR-2004.

29-SEP-2003; 2003WO-US028547.

02-OCT-2002; 2002US-0414971P.

(GETH) GENENTECH INC.

Wu TD, Zhang Z, Zhou Y;

WPI; 2004-347921/32.

P-PSDB; AEM80804.

New tumor-associated antigenic target polypeptides and nucleic acids, useful in preparing a medicament for treating or detecting a proliferative disorder, e.g. breast, lung, colorectal, ovarian or prostate cancer or tumor.

Claim 1; SEQ ID NO 2070; 7273pp; English.

The invention relates to human tumour-associated antigenic target (TAT) polypeptides, and their related nucleic acids. The TAT polypeptides are overexpressed in cancer tissues compared to normal tissues, and may thus serve as effective targets for the diagnosis and treatment of cancer in mammals. The invention also relates to nucleic acid and polypeptide sequences at least 80% identical to the TAT nucleic acids and polypeptides; expression vectors and host cells comprising a TAT nucleic acid; an antibody specific for a TAT polypeptide; a peptide or organic molecule which binds to a TAT polypeptide; fusion proteins comprising a TAT polypeptide; and methods and compositions for the treatment or diagnosis of cancer in mammals. TAT polypeptides, nucleic acids, antibodies, antagonists, binding molecules and compositions are useful for diagnosing or treating a cell proliferative disorder associated with increased TAT expression, particularly cancers such as breast cancer, colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder cancer, pancreatic cancer, cervical cancer, cancers of the central nervous system, melanoma and leukaemia. TAT nucleic acids may further be used as hybridisation probes, in chromosome and gene mapping, in chromosome identification and in gene therapy. The present sequence represents a TAT nucleic acid of the invention

Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	19.2	Length:	2053
Score:	48.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	13	Gaps:	0

US-10-774-176-9 (1-9) x ACN38510 (1-2053)

QY 1 GlyLeuProHisIleArgValPheLeu 9

DB 931 GGTCTACCCACATTAGGTTTCTG 957

RESULT 23

ADV35098

ID ADV35098 standard; cDNA; 2053 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

AC ADV35098;
XX 10-FEB-2005 (first entry)
XX Human cDNA of an exemplary efficacy gene for BAD SeqId174.
XX human; ss; multi-parameter high throughput screening; MPHTS;
KW disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
KW bipolar affective disorder; BAD; autism; Parkinson's;
KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.
XX
XX Homo sapiens.
XX US2003096264-A1.
XX 22-MAY-2003.
XX 18-JUN-2002; 2002US-00175523.
XX 18-JUN-2001; 2001US-0299151P.
XX 07-SEP-2001; 2001US-0317828P.
XX 25-SEP-2001; 2001US-0325150P.
XX 14-NOV-2001; 2001US-0333047P.
XX 18-JAN-2002; 2002US-0349936P.
XX 04-MAR-2002; 2002US-0361834P.
XX (PSYC-) PSYCHIATRIC GENOMICS INC.
XX
XX Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
XX Palfreyman M, Rajan P;
XX WPI; 2004-118903/12.
XX Identifying a compound that can treat disease or disorders, such as, a
XX neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
XX determining the expression of one or more efficacy genes in a cell
XX contacted with the test compound.
XX
XX Example 6; SEQ ID NO 174; 39pp; English.
XX
XX This invention relates to a novel screening method identified as a multi-
XX parameter high throughput screening (MPHTS) assay. Specifically, it
XX refers to an assay that utilizes the disease signature of a plurality of
XX specific genes associated with a particular disease, and identifies
XX differential expression between those cells taken from individuals
XX affected by that disease and those that are not affected. The present
XX invention then describes the screening of candidate pharmaceutical
XX compounds to identify those that have a potential therapeutic benefit for
XX the treatment of neuropsychiatric and neurodegenerative disorders
XX including schizophrenia, bipolar affective disorder (BAD) and autism, as
XX well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
XX of this invention exhibit various activities including neuroleptic,
XX nootropic, antimanic and antidepressant. Furthermore, the screening
XX method used in MPHTS will be automated, such that a large number of test
XX compounds may be rapidly screened with a minimal amount of labour and
XX effort. This polynucleotide is a human cDNA sequence of a gene that is
XX differentially expressed in the presence of a therapeutic compound and
XX represents an exemplary efficacy gene for bipolar affective disorder,
XX given in an exemplification of the invention.
XX
XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 19.2 Length: 2053
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 13 Indels: 0
DB: Gaps: 0

US-10-774-176-9 (1-9) x ADV35098 (1-2053)
Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 931 GGTCTACCCCACTTAGGGTTTCCTG 957
RESULT 24
AAS87175
ID AAS87175 standard; cDNA; 2338 BP.
XX AAS87175;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #22979.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG223988.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 22979; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 22.3 Length: 2338
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 5

US-10-774-176-9 (1-9) x AAS87175 (1-2338)

QY 1 GlyLeuProHisIleArgValPheLeu 9
 DB 1188 GGTCTACCCACATTAGGGTTTCCTG 1214

RESULT 25

AAK94253
 ID AAK94253 standard; cDNA; 2359 BP.

XX AAK94253;

XX 06-NOV-2001 (first entry)

XX Human full-length cDNA, SEQ ID NO: 2864.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI: 2001-524255/58.

DR P-PSDB; AAK93333.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

XX in genetic manipulation.

XX Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesizing full length cDNA

XX clones. 830 cDNA molecules encoding a human protein have been isolated

XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have

XX been determined. Primers for synthesizing the full length cDNA are useful

XX for clarifying the function of the protein encoded by the cDNA. The full

XX length clones were obtained by construction of full length enriched cDNA

XX libraries that were synthesised by the oligo-capping method. The primers

XX enable the production of the full length cDNA easily without any special

XX methods. The present sequence is a full length human cDNA of the

XX invention. Note: The sequence data for this patent did not form part of

XX the printed specification, but was obtained in CD-ROM format directly

XX from EPO

XX Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 22.5 Length: 2359

Score: 48.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 4 Gaps: 0

US-10-774-176-9 (1-9) x AAK94253 (1-2359)

QY 1 GlyLeuProHisIleArgValPheLeu 9

DB 1270 GGTCTACCCACATTAGGGTTTCCTG 1296

RESULT 26

ADL30831

ID ADL30831 standard; cDNA; 2359 BP.

XX ADL30831;

XX 20-MAY-2004 (first entry)

XX Full length human cDNA clone SeqID 2864.

XX human; medicine; signal transduction; glycoprotein; transcription;

XX oligo-capping method; ss; gene.

XX Homo sapiens.

XX EP1396543-A2.

XX 10-MAR-2004.

XX 07-JUL-2000; 2003EP-00025638.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183865.

XX 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ota T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI: 2004-204755/20.

XX P-PSDB; ADL30832.

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full

XX length human cDNAs.

XX Example 1; SEQ ID NO 2864; 1340pp; English.

XX This invention relates to a novel primers useful for synthesizing full

XX length cDNA molecules that encode human proteins. Specifically, it refers

XX to secretory or membrane proteins that are potential therapeutic agents/

XX target molecules in the field of medicine, and in particular genes

XX encoding proteins that are associated with signal transduction,

XX glycoproteins and transcription. The present invention describes a method

XX for efficiently cloning a full length human cDNA from both the 5' and 3'

XX ends using the oligo-capping method. This polynucleotide sequence is a

XX full length human cDNA clone of the invention.

XX Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 22.5 Length: 2359

Score: 48.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 12 Gaps: 0

US-10-774-176-9 (1-9) x ADL30831 (1-2359)

QY 1 GlyLeuProHisIleArgValPheLeu 9

DB 1270 GGTCTACCCACATTAGGGTTTCCTG 1296

RESULT 27

AAK94254

ID AAK94254 standard; cDNA; 2361 BP.

XX AAK94254;

XX 06-NOV-2001 (first entry)

XX Human full-length cDNA, SEQ ID NO: 2866.

```

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-00114089.
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX (HELI-) HELIX RES INST.
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX P-PSDB; AAM93334.
XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX in genetic manipulation.
XX Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.
XX The invention relates to primers for synthesising full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesising the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a full length human cDNA of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in CD-ROM format directly
XX from BPO
XX
XX Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 22.6 Length: 2361
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-9 (1-9) x AAK94254 (1-2361)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 1272 GGTCTACCCACATTAGGGTTTCCTG 1298

RESULT 28
AD126162
ID AD126162 standard; cDNA; 2361 BP.
XX
AC AD126162;
XX
XX 22-APR-2004 (first entry)
XX Human cDNA encoding protein that promotes STAT6 activation #64.
XX
XX ss; gene; human; signal transducer and activator of transcription 6;
XX STAT6; immunogen; STAT6 activation; allergy; inflammation;
XX autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
XX Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
XX systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
XX ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

```

```

XX Homo sapiens.
XX WO2003104277-A2.
XX 18-DEC-2003.
XX 05-JUN-2003; 2003WO-JP007123.
XX 05-JUN-2002; 2002JP-00164257.
XX 06-JUN-2002; 2002US-0385912P.
XX 26-DEC-2002; 2002JP-00377326.
XX 27-DEC-2002; 2002US-0436467P.
XX 15-MAY-2003; 2003JP-00137505.
XX 16-MAY-2003; 2003US-0470836P.
XX (ASAH) ASahi KASEI KK.
XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
XX WPI; 2004-122214/12.
XX P-PSDB; ADI26163.
XX
XX New signal transducer and activator of transcription 6 activation
XX promoting purified protein, for diagnosing and treating disease
XX associated with activation/inhibition of transcription factor e.g.
XX diabetes and cancer.
XX
XX Claim 4; SEQ ID NO 127; 1368pp; English.
XX The invention relates to a purified protein promoting signal transducer
XX and activator of transcription 6 activation (STAT6). The protein is
XX useful for the producing an antibody, which involves administering the
XX protein or its epitope-bearing fragments to a non-human animal as an
XX antigen. The nucleic acid is useful for diagnosing a disease or
XX susceptibility to a disease related to the protein or activity of the
XX protein. A transformant expressing the protein is useful for screening
XX compounds which inhibit or promote STAT6 activation. A transformant
XX expressing the protein is useful for producing a pharmaceutical
XX composition. Compositions, antibodies and antisense molecules are useful
XX for the treating a disease associated with STAT6 activation such as
XX allergic diseases, inflammation, autoimmune diseases, diabetes,
XX hyperlipidaemia, infectious disease and cancers. Compositions are useful
XX for treating disease associated with STAT6 activation and/or prevention
XX of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
XX arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
XX allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
XX viral hepatitis and AIDS. The protein has efficiently promoting STAT6
XX activity. The protein or nucleic acid is effectively useful for screening
XX compounds for treating and preventing disease associated with excessive
XX activation or inhibition of STAT6. The present sequence represents a
XX human cDNA encoding a protein which promotes STAT6 activation.
XX
XX Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 22.6 Length: 2361
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-9 (1-9) x ADI26162 (1-2361)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 1272 GGTCTACCCACATTAGGGTTTCCTG 1298

RESULT 29
AD130833
ID ADL30833 standard; cDNA; 2361 BP.
XX

```

AC ADL30833;
 XX 20-MAY-2004 (first entry)
 XX Full length human cDNA clone SeqID 2866.
 XX human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.
 KW Homo sapiens.
 OS
 XX EP1396543-A2.
 XX 10-MAR-2004.
 XX 07-JUL-2000; 2003EP-00025638.
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2004-204755/20.
 DR P-PSDB; ADL30834.
 XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX Example 1; SEQ ID NO 2866; 1340pp; English.
 XX This invention relates to a novel primers useful for synthesising full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.
 XX Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. NO.: 22.6 Length: 2361
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-9 (1-9) x ADL30833 (1-2361)
 Qy 1 GlyLeuProHisIleArgValPheLeu 9
 Db 1272 GGTCTACCCACATTAGGGTTTCCTG 1298
 RESULT 30
 ADL30833
 ID ADS96565 standard; cDNA; 1587 BP.
 XX
 AC ADS96565;
 XX
 DT 02-DEC-2004 (first entry)
 XX Drosophila melanogaster protein coding sequence, SEQ ID 186.
 DE Insecticide; Antiparasitic; Antihelminthic; gene; ds.
 KW
 XX

OS Drosophila melanogaster.
 XX WO2004039999-A2.
 XX 13-MAY-2004.
 XX 08-AUG-2003; 2003WO-US024982.
 XX 30-OCT-2002; 2002US-0422377P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA Stam L, Kamdar KP, Spana E, Bachmann J;
 PI WPI; 2004-376203/35.
 DR P-PSDB; ADS96566.
 XX Identifying a compound that inhibits the activity of a protein for
 PT Drosophila viability for use e.g., as insecticidal agent by expressing in
 PT a recombinant host a DNA molecule to produce a protein essential for
 PT Drosophila viability.
 XX Claim 1; SEQ ID NO 186; 57pp; English.
 PS The present invention relates to a method for identifying a compound that
 CC inhibits the activity of a protein essential for Drosophila viability.
 CC The method comprises: (a) expressing in a recombinant host a DNA sequence
 CC encoding a protein essential for Drosophila viability; (b) testing
 CC compounds suspected of having the ability to inhibit the activity of the
 CC protein expressed in (a); and identifying a compound tested in (b) that
 CC inhibits the activity of the protein. The method is useful in identifying
 CC a compound that inhibits the activity of a protein essential for
 CC Drosophila viability for use as insecticidal, ectoparasiticide,
 CC antiparasitic, antihelminthic or acaricidal agent. The present sequence
 CC is the DNA sequence for one such protein essential for Drosophila
 CC viability. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1587 BP; 394 A; 432 C; 383 G; 378 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. NO.: 230 Length: 1587
 Score: 42.00 Matches: 6
 Percent Similarity: 100.0% Conservative: 3
 Best Local Similarity: 66.7% Mismatches: 0
 Query Match: 87.5% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-9 (1-9) x ADS96565 (1-1587)
 Qy 1 GlyLeuProHisIleArgValPheLeu 9
 Db 603 GGCCCTTCCTCCACCTACGATATATCTT 629
 RESULT 31
 ABL20593
 ID ABL20593 standard; DNA; 1704 BP.
 XX
 AC ABL20593;
 XX
 DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 13252.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 KW Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 XX 27-SEP-2001.
 PD

XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 13252; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1704 BP; 428 A; 457 C; 404 G; 415 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 250 Length: 1704
Score: 42.00 Matches: 6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match: 87.5% Indels: 0
DB: 4 Gaps: 0
US-10-774-176-9 (1-9) x ABL20593 (1-1704)
QY 1 GlyLeuProHisIleArgValPheLeu 9
DB 720 GGCCTTCCTCACCACGATATATCTT 746
RESULT 32
ABL20592
ID ABL20592 standard; DNA; 4357 BP.
AC ABL20592;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 13249.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX

DR WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 13249; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 4357 BP; 1274 A; 932 C; 855 G; 1296 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 747 Length: 4357
Score: 42.00 Matches: 6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match: 87.5% Indels: 0
DB: 4 Gaps: 0
US-10-774-176-9 (1-9) x ABL20592 (1-4357)
QY 1 GlyLeuProHisIleArgValPheLeu 9
DB 2373 GGCCTTCCTCACCACGATATATCTT 2399
RESULT 33
ACN80343/C
ID ACN80343 standard; DNA; 787 BP.
XX
XX ACN80343;
XX
XX 02-DEC-2004 (first entry)
XX
XX Breast cancer related marker, seq id 1493.
XX
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX Homo sapiens.
XX
XX US2003099974-A1.
XX
XX 29-MAY-2003.
XX
XX 18-JUL-2002; 2002US-00198846.
XX
XX 18-JUL-2001; 2001US-0306220P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2003-787014/74.
XX
XX Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
XX Disclosure; SEQ ID NO 1493; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the

US-10-774-176-9 (1-9) x AAA27060 (1-901)
 QY 2 LeuProHisIleArgValPheLeu 9
 DB 365 CTGCCCCAGTCCGGGCTCTCTCTG 388

RESULT 35
 ABK87175
 ID ABK87175 standard; cDNA; 1260 BP.

XX AC ABK87175;

XX DT 07-OCT-2002 (first entry)

XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.

XX Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.

XX OS Felis sp.

XX FH Key Location/Qualifiers
 XX CDS 1..1260

XX FT /*tag= a

XX FT /product= "5T4 protein"

XX PN WO200238612-A2.

XX PD 16-MAY-2002.

XX PF 13-NOV-2001; 2001WO-GB005004.

XX PR 13-NOV-2000; 2000WO-GB004317.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI Myers K, Drury N, Carroll M;

XX DR WPI; 2002-557449/59.

XX DR P-PSDB; AAU98694.

XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.

XX Claim 4; Page 68; 68pp; English.

XX The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes feline 5T4 protein

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 280 Length: 1260

Score: 41.00 Matches: 7

Percent Similarity: 100.0% Conservative: 1

Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 85.4% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-9 (1-9) x ABK87175 (1-1260)

QY 2 LeuProHisIleArgValPheLeu 9

DB 847 CTGCCCCAGTCCGGGCTCTCTCTG 870

RESULT 36

ADB97513

ID ADB97513 standard; DNA; 1260 BP.

XX AC ADB97513;

XX DT 04-DEC-2003 (first entry)

XX DE Feline 5T4 antigen DNA.

XX KW Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX CDS 1..1260

XX FT /*tag= a

XX FT /product= "Feline 5T4 antigen protein"

XX PN WO2003068816-A1.

XX PD 21-AUG-2003.

XX PF 13-FEB-2003; 2003WO-GB000670.

XX PR 13-FEB-2002; 2002GB-00003419.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI Carroll M, Kingsman S, Redchenko I;

XX DR WPI; 2003-637141/60.

XX DR P-PSDB; ADB97520.

XX New major histocompatibility complex class I peptide epitopes from human
 PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.

XX Disclosure; Page 67; 73pp; English.

XX The invention relates to a novel Major Histocompatibility Complex (MHC)
 CC class I peptide epitope of the 5T4 antigen. The invention further
 CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
 CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
 CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 5T4 epitope and/or its encoding nucleic acid; a method
 CC comprising detecting the presence of the 5T4 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 5T4 epitope in conjunction with an MHC class
 CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC particularly cancer. The detection method is useful for diagnosing or
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline

CC 5T4 antigen coding DNA of the invention.

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 280 Length: 1260

Score: 41.00 Matches: 7

Percent Similarity: 100.0% Conservatives: 1

Best Local Similarity: 87.5% Mismatches: 0

Query Match: 85.4% Indels: 0

DB: 10 Gaps: 0

US-10-774-176-9 (1-9) x ADB97513 (1-1260)

QY 2 LeuProHisIleArgValPheLeu 9

DB 847 CTGCCCCACGTCAGGCTCTTCCTG 870

RESULT 37

ADB97452

ID ADB97452 standard; DNA; 1260 BP.

XX AC ADB97452;

DT 04-DEC-2003 (first entry)

DE DNA encoding feline 5T4 protein.

XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;

KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT CDS 1..1260

FT /*tag= a

FT /product= "Feline 5T4 antigen protein"

XX WO20003068815-A2.

XX PD 21-AUG-2003.

XX PF 13-FEB-2003; 2003WO-GB0000618.

XX PR 13-FEB-2002; 2002GB-00003420.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI Carroll M, Harrop R, Kingsman S;

XX WPI; 2003-663795/62.

XX P-PSDB; ADB97455.

PT New Major Histocompatibility Complex class II peptide epitope of 5T4,

PT useful for manufacturing a medicament for diagnosing, preventing and/or

PT treating a disease, e.g. cancer.

XX PS Disclosure; Page 49; 63pp; English.

XX CC The invention relates to a Major Histocompatibility Complex (MHC) class

CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or

CC clone has a cytostatic activity, as it is useful in manufacturing a

CC medicament for preventing and/or treating a disease, particularly cancer.

CC The methods are useful for detecting T-cells capable of specifically

CC recognising a peptide epitope in conjunction with an MHC molecule, for

CC diagnosing or monitoring the progression of a cancerous disease, or for

CC detecting the presence of a peptide or nucleic acid using an agent. The

CC MHC class II peptide epitope of the invention can be used in gene therapy

CC or as part of a vaccine. This polynucleotide sequence represents the DNA

CC coding for the feline 5T4 protein.

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 280 Length: 1260

Score: 41.00 Matches: 7

Percent Similarity: 100.0% Conservatives: 1

Best Local Similarity: 87.5% Mismatches: 0

Query Match: 85.4% Indels: 0

DB: 10 Gaps: 0

US-10-774-176-9 (1-9) x ADB97452 (1-1260)

QY 2 LeuProHisIleArgValPheLeu 9

DB 847 CTGCCCCACGTCAGGCTCTTCCTG 870

RESULT 38

AAF89736

ID AAF89736 standard; DNA; 1263 BP.

XX AC AAF89736;

DT 23-JUL-2001 (first entry)

DE Nucleotide sequence of canine 5T4 protein.

XX KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;

KW hypersensitivity; autoimmune disease; central nervous system disease;

KW Parkinson's disease; periodontal disease; cardiopulmonary disease;

KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;

KW Helicobacter-related disease; immune disorder; ss.

XX OS Canis sp.

XX FH Key Location/Qualifiers

FT CDS 1..1263

FT /*tag= a

FT /product= "5T4"

XX WO200136486-A2.

XX PD 25-MAY-2001.

XX PF 13-NOV-2000; 2000WO-GB004317.

XX PR 18-NOV-1999; 99WO-GB003859.

PR 15-FEB-2000; 2000GB-00003527.

PR 02-MAR-2000; 2000GB-00005071.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;

PI Myers KA;

XX WPI; 2001-343805/36.

XX P-PSDB; AAB83839.

PT Use of single chain antibody capable of recognizing a disease associated

PT molecule for manufacturing a medicament for preventing and/or treating a

PT disease condition associated with disease associated molecule.

XX PS Disclosure; Fig 26; 118pp; English.

XX CC The specification describes the use of a single chain antibody (ScFv),

CC which is capable of recognizing a disease associated molecule in the

CC manufacture of a medicament for the prevention and treatment of a disease

CC condition. The ScFv antibody is useful in the manufacture of a

CC medicament, for affecting a disease in vivo, for preparing a

CC pharmaceutical composition, for in vivo imaging and/or for adjuvant

CC treatment of a disease. The ScFv antibody is also useful for treating

CC inflammatory diseases including arthritis, hypersensitivity, autoimmune

CC diseases, cancers, central nervous system disorders including Parkinson's

CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular

CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-

CC related diseases, and other immune disorders. The present sequence

CC encodes a 5T4 protein, which is used to produce scFv of the invention
XX Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 281 Length: 1263
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 85.4% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-9 (1-9) x AAF89736 (1-1263)

Qy 2 LeuProHisIleArgValPheLeu 9
|||||:|||||
Db 850 CTGCCCCAGTCCGGTCTCTCTG 873

RESULT 39

ABK87174

ID ABK87174 standard; cDNA; 1263 BP.

XX AC ABK87174;

XX DT 07-OCT-2002 (first entry)

XX DE cDNA encoding canine oncofetal leucine-rich glycoprotein, 5T4.

XX KW Canine; dog; oncofetal leucine-rich glycoprotein; 5T4; tumour;
XX cell proliferative disorder; infection; inflammatory condition;
XX cancer immunotherapy; foetal cell; maternal blood; cytostatic;
XX foetal abnormality; foetal sex determination; gene; sp.

XX OS Canis sp.

XX FH Key Location/Qualifiers
XX CDS 1..1263
XX FT /*tag= a
XX FT /product= "5T4 protein"

XX PN WO200238612-A2.

XX PD 16-MAY-2002.

XX PF 13-NOV-2001; 2001MO-GB005004.

XX PR 13-NOV-2000; 2000MO-GB004317.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI Myers K, Drury N, Carroll M;

XX PT WPI; 2002-557449/59.

XX DR P-PSDB; AAU98693.

XX FT Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
XX polypeptide, useful in preparation of vaccine for treating and/or
XX preventing cancer in a subject, preferably a dog or cat.

XX PS Claim 1; Page 67; 68pp; English.

XX CC The present invention relates to the isolation of canine and feline
XX oncofetal leucine-rich glycoproteins known as 5T4, and the
XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
XX a significant proportion of tumours. The sequences of the invention are
XX useful in a pharmaceutical composition for the prevention and/or
XX treatment of tumours or other diseases associated with cell
XX proliferation, infections, and inflammatory conditions in animals,
XX preferably dogs or cats. The compositions may also be used for cancer
XX immunotherapy in these animals. The sequences of the invention may also
XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
XX measurement and localisation of 5T4 in extracts of plasma, urine,
XX tissues, and in cell culture media. Antibodies specific for the 5T4

CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus(es).
XX The present sequence encodes canine 5T4 protein

SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 281 Length: 1263
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 85.4% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-9 (1-9) x ABK87174 (1-1263)

Qy 2 LeuProHisIleArgValPheLeu 9
|||||:|||||

Db 850 CTGCCCCAGTCCGGTCTCTCTG 873

RESULT 40

ADO34435_3

Continuation (4 of 7) of ADO34435 from base 300001 (Human SLIT-3 genomic sequence.)
WP Sequence split into 7 fragments LOCUS ADO34435 Accession ADO34435

WP	Fragment Name	Begin	End
WP	ADO34435_0	1	110000
WP	ADO34435_1	100001	210000
WP	ADO34435_2	200001	310000
WP	ADO34435_3	300001	410000
WP	ADO34435_4	400001	510000
WP	ADO34435_5	500001	610000
WP	ADO34435_6	600001	634886

Alignment Scores:

Pred. No.: 5.11e+04 Length: 110000
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 85.4% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-9 (1-9) x ADO34435_3 (1-110000)

Qy 2 LeuProHisIleArgValPheLeu 9
|||||:|||||

Db 103442 TTGCCTCACATCCGTATCTTCTC 103465

RESULT 41

ADO34435_4

Continuation (5 of 7) of ADO34435 from base 400001 (Human SLIT-3 genomic sequence.)
WP Sequence split into 7 fragments LOCUS ADO34435 Accession ADO34435

WP	Fragment Name	Begin	End
WP	ADO34435_0	1	110000
WP	ADO34435_1	100001	210000
WP	ADO34435_2	200001	310000
WP	ADO34435_3	300001	410000
WP	ADO34435_4	400001	510000
WP	ADO34435_5	500001	610000
WP	ADO34435_6	600001	634886

Alignment Scores:

Pred. No.: 5.11e+04 Length: 110000
Score: 41.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 85.4% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-9 (1-9) x ADO34435_4 (1-110000)

Qy 2 LeuProHisIleArgValPheLeu 9

Db 3442 TTGCTTCACATCGTATCTTCTC 3465
 RESULT 42
 ID ABX36840 standard; cDNA; 416 BP.
 XX AC ABX36840;
 XX DT 20-FEB-2003 (first entry)
 XX DE Bovine EST associated with lactation/muscle/fat deposition #2005.
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 XX KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX KW gene analysis; cattle breeding.
 XX OS Bos Taurus.
 XX PN US2002137139-A1.
 XX PD 26-SEP-2002.
 XX PF 24-SEP-2001; 2001US-00960352.
 XX PR 12-JAN-1999; 99US-0115707P.
 XX PR 11-JAN-2000; 2000US-00480902.
 XX PA (BYAT/) BYATT J C.
 XX PA (MATH/) MATHIALAGAN N.
 XX PA (TAON/) TAO N.
 XX PA (WARR/) WARREN W C.
 XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX Claim 2; SEQ ID NO 2005; 245pp; English.
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived from
 CC cattle, and the LMPD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
 CC acid linked to a promoter and a 3' non- translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMPD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, or
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX SQ Sequence 416 BP; 110 A; 60 C; 75 G; 171 T; 0 U; 0 Other;
 Alignment Scores:

Pred. No.: 122 Length: 416
 Score: 40.00 Matches: 8
 Percent Similarity: 88.9% Conservative: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 83.3% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-9 (1-9) x ABX36840 (1-416)
 QY 1 GlyLeuProHisIleArgValPheLeu 9
 Db 287 GGCTTACCTCACATTTTGTGTTTCTC 313
 RESULT 43
 ID ABX49697 standard; cDNA; 433 BP.
 XX AC ABX49697;
 XX DT 21-FEB-2003 (first entry)
 XX DE Bovine EST associated with lactation/muscle/fat deposition #14862.
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 XX KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX KW gene analysis; cattle breeding.
 XX OS Bos Taurus.
 XX PN US2002137139-A1.
 XX PD 26-SEP-2002.
 XX PF 24-SEP-2001; 2001US-00960352.
 XX PR 12-JAN-1999; 99US-0115707P.
 XX PR 11-JAN-2000; 2000US-00480902.
 XX PA (BYAT/) BYATT J C.
 XX PA (MATH/) MATHIALAGAN N.
 XX PA (TAON/) TAO N.
 XX PA (WARR/) WARREN W C.
 XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX Claim 2; SEQ ID NO 14862; 245pp; English.
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived from
 CC cattle, and the LMPD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
 CC acid linked to a promoter and a 3' non- translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMPD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle

CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139

XX
SQ Sequence 433 BP; 121 A; 75 C; 77 G; 159 T; 0 U; 1 Other;

Alignment Scores: Length: 433
Pred. No.: 128 Matches: 8
Score: 40.00
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 83.3% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x ABX49697 (1-433)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 167 GGGTACCTGCACATTTTGTGTTCTC 193

RESULT 44

AAH36486/C

ID AAH36486 standard; cDNA; 465 BP.

XX AC AAH36486;

XX DT 03-SEP-2001 (first entry)

XX DE Human colon cancer antigen encoding cDNA SEQ ID NO:3568.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; ss.

XX OS Homo sapiens.

XX PN WO200122920-A2.

XX PD 05-APR-2001.

XX PF 28-SEP-2000; 2000WO-US026524.

XX PR 29-SEP-1999; 99US-0157137P.

XX PR 03-NOV-1999; 99US-0163280P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX P-PSDB; AAG77081.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 1; Page 5353-5354; 9803pp; English.

XX AAH32943 to AAH37195 and AAG77195 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where the
XX proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene therapy
XX and vaccine production. N and P may be used in the prevention, diagnosis
XX and treatment of diseases associated with inappropriate P expression. For
XX example, N and P may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of P by expressing inactive proteins or to
XX supplement the patient's own production of P. Additionally, N may be used
XX to produce the colon cancer-associated P, by inserting the nucleic acids
XX into a host cell and culturing the cell to express the proteins. N and P
XX can be used in the prevention, diagnosis and treatment of colorectal
XX carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent

CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922

XX SQ Sequence 465 BP; 96 A; 149 C; 147 G; 68 T; 0 U; 5 Other;

Alignment Scores: Length: 465
Pred. No.: 139 Matches: 7
Score: 40.00
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 83.3% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-9 (1-9) x AAH36486 (1-465)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 354 GGGCTACCCCATTTGAGGCTCCTGCTT 328

RESULT 45

ACH45478/C

ID ACH45478 standard; cDNA; 467 BP.

XX AC ACH45478;

XX DT 13-OCT-2003 (first entry)

XX DE Human foetal brain cDNA #6203.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX genome mapping; biodiversity; genetic disorder.

XX OS Homo sapiens.

XX PN US2003073623-A1.

XX PD 17-APR-2003.

XX PF 30-JUL-2001; 2001US-00918995.

XX PR 30-JUL-2001; 2001US-00918995.

XX PA (DRMA/) DRMANAC R T.

XX PA (LABA/) LABAT I.

XX PA (STAC/) STACHE-CRAIN B.

XX PA (DICK/) DICKSON M C.

XX PA (JONE/) JONES L W.

XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 32690; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide

PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 11270-11271; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 565 BP; 164 A; 146 C; 115 G; 137 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 175 Length: 565
Score: 40.00 Matches: 7
Percent Similarity: 87.5% Conservative: 0
Best Local Similarity: 87.5% Mismatches: 1
Query Match: 83.3% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-9 (1-9) x ABV58745 (1-565)

QY 1 GlyLeuProHisIleArgValphe 8
|||||
DB 297 GGTTCCTCATTCAGAGTGT 320

RESULT 48

ABV56637
ID ABV56637 standard; cDNA; 615 BP.

AC ABV56637;

XX 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 56628.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JB;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 10920-10921; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 615 BP; 164 A; 155 C; 127 G; 166 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 193 Length: 615
Score: 40.00 Matches: 7
Percent Similarity: 87.5% Conservative: 0
Best Local Similarity: 87.5% Mismatches: 1
Query Match: 83.3% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-9 (1-9) x ABV56637 (1-615)

QY 1 GlyLeuProHisIleArgValphe 8

DB 302 GGTTCCTCATTCAGAGTGT 325

RESULT 49

AAH52150/c

ID AAH52150 standard; cDNA; 933 BP.

XX AAH52150;

XX 10-SEP-2001 (first entry)

XX Human AFP protein encoding cDNA sequence SEQ ID NO:115.

XX Human; secreted protein; secretion; bacterial cell; fungal cell; eukaryotic cell; fusion protein; maltose binding protein; immunoglobulin constant region; polyhistidine tag; ss.

XX Homo sapiens.

XX WO200129221-A2.

XX 26-APR-2001.

XX 20-OCT-2000; 2000WO-US029052.

XX 20-OCT-1999; 99US-0160712P.

XX (ZYMO) ZYMOGENETICS INC.

XX Conklin DC, Yee DP;

XX WPI; 2001-300340/31.

XX P-P5DB; AAG81299.

XX Isolated polypeptide for directing secretion of proteins of interest from a host cell including, e.g. bacteria, includes contiguous amino acid residues of polypeptide with specified amino acids.

XX Claim 9; Page 231-233; 617pp; English.

XX AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242 to AAG81453. The secreted proteins can be used for directing the secretion of proteins of interest from a host cell including bacteria, fungal cells, and cultured higher eukaryotic cells. The present invention also describes fusion proteins, where a secreted protein of the invention

CC is operably linked via a peptide bond or peptide linker to a second
 CC protein selected from the group consisting of maltose binding protein, an
 CC immunoglobulin constant region, a polystyrene tag and a peptide given
 CC in AAG81453

SQ Sequence 933 BP; 163 A; 311 C; 302 G; 156 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	314	Length:	933
Score:	40.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	83.3%	Indels:	0
DB:	5	Gaps:	0

US-10-774-176-9 (1-9) x AAHS2150 (1-933)

QY 1 GlyLeuProHisIleArgValPheLeu 9

DB 529 GGGTACCCCAATTGAGGGTCTGCTT 503

RESULT 50

ID AQ87043/c

XX AQ87043 standard; cDNA; 1333 BP.

AC AQ87043;

XX 07-OCT-2004 (first entry)

DT Human tumour-associated antigenic target (TAT) cDNA sequence #3919.

DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

KW cancer; cell proliferative disorder; gene; ss.

XX Homo sapiens.

XX WO2004060270-A2.

XX 22-JUL-2004.

PF 15-OCT-2003; 2003WO-US029126.

XX 18-OCT-2002; 2002US-0418988P.

XX (GETH) GENENTECH INC.

PA (WUTD/) WU T D.

PA (ZHOU/) ZHOU Y.

XX Wu TD, Zhou Y;

XX WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 3919; 5504pp; English.

XX The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)

CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.

XX
 SQ Sequence 1333 BP; 238 A; 397 C; 413 G; 285 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	475	Length:	1333
Score:	40.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	83.3%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-9 (1-9) x AQ87043 (1-1333)

QY 1 GlyLeuProHisIleArgValPheLeu 9

DB 625 GGGTACCCCAATTGAGGGTCTGCTT 599

Search completed: April 25, 2006, 12:38:19

Job time : 327.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search using frame plus p2n model

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds

(without alignments)
171.290 Million cell updates/sec

Title: US-10-774-176-9

Perfect score: 48

Sequence: 1 GLPHRVPL 9

Scoring table:

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abes/ABSSWEB spool/US10774176/runat_24042006.165114.19197/app_query.fasta_1
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes04
-USER=US10774176 @GCN 1.1 6765 @runat_24042006.165114.19197 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.st.*

11: gb.sv.*

12: gb.un.*

13: gb.vi.*

14: gb.htg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	475	6	CQ920916 Sequence
2	48	100.0	927	6	AX829164 Sequence
3	48	100.0	1263	6	BD249731 Polypepti

4	48	100.0	1263	6	AX025011
5	48	100.0	1263	6	AX316086
6	48	100.0	2053	6	CQ731678
7	48	100.0	2053	8	HS5740A
8	48	100.0	2359	6	BD127282
9	48	100.0	2359	8	CQ782724
10	48	100.0	2359	8	AK074786
11	48	100.0	2361	6	BD127283
12	48	100.0	2361	6	CQ782726
13	48	100.0	2361	6	AX961916
14	48	100.0	2361	8	AK074790
15	48	100.0	2379	8	BC037161
16	48	100.0	5551	8	HS04012159
17	48	100.0	121309	8	HS0492P14
18	47	97.9	2714	8	AB168308
19	44	91.7	152537	8	AC055731
20	44	91.7	160094	5	EX855617
21	44	91.7	191959	14	AC012217
22	44	91.7	246716	14	AC096699
23	43	89.6	779	10	BV595106
24	43	89.6	24950	2	AF014939
25	43	89.6	72957	14	AC025304
26	43	89.6	89743	9	AC003062
27	43	89.6	110000	1	AE013598_20
28	43	89.6	136897	9	AC005816
29	43	89.6	147412	9	AC079831
30	43	89.6	153264	14	AC135216
31	43	89.6	158546	14	AC022876
32	43	89.6	174699	8	AC009949
33	43	89.6	183066	14	AC123238
34	43	89.6	184044	9	AC008020
35	43	89.6	202607	8	AP002353
36	43	89.6	204343	14	AC102169
37	43	89.6	212065	14	AC166253
38	43	89.6	215042	14	AC162395
39	43	89.6	221701	14	AC133670
40	43	89.6	222589	9	AC133284
41	43	89.6	224138	9	AC133573
42	43	89.6	231951	14	AC158595
43	43	89.6	232670	14	AC129149
44	43	89.6	252441	14	AC099417
45	43	89.6	258569	14	AC097908
46	43	89.6	263882	14	AC122618
47	42	87.5	563	6	AR520225
48	42	87.5	1017	6	AR504928
49	42	87.5	1704	6	CQ600372
50	42	87.5	1980	2	BT006007
51	42	87.5	1987	2	AY060891
52	42	87.5	4357	6	CQ600371
53	42	87.5	24145	14	AC014455
54	42	87.5	31110	2	CBRG14K24
55	42	87.5	110000	1	BA000040_22
56	42	87.5	110000	14	AC118422_2
57	42	87.5	110000	15	AP008217_220
58	42	87.5	130492	8	AC006357
59	42	87.5	156069	15	AC136842
60	42	87.5	157708	8	AP005357
61	42	87.5	163546	14	AC103547
62	42	87.5	164386	2	AC006491
63	42	87.5	170298	2	AB003690
64	42	87.5	180496	14	AC117328
65	42	87.5	188776	9	BX537258
66	42	87.5	190151	14	AC149113
67	42	87.5	203449	9	AL713960
68	42	87.5	210536	14	AC165158
69	42	87.5	210842	14	AC140748
70	42	87.5	211108	14	AC129741
71	42	87.5	212999	14	AC111972
72	42	87.5	213495	9	AL831778
73	42	87.5	213660	14	AC073699
74	42	87.5	218672	14	AC103267
75	42	87.5	220661	14	AC094122
76	42	87.5	223642	14	AC108290


```

c 953 38 79.2 263827 14 AC094579 AC094579 Rattus no
c 954 38 79.2 264628 14 AC121690 AC121690 Rattus no
c 955 38 79.2 265393 14 CR974457 CR974457 Mus muscu
c 956 38 79.2 265802 14 AC118507 AC118507 Rattus no
c 957 38 79.2 265807 14 AC130507 AC130507 Rattus no
c 958 38 79.2 267708 9 AC113320 AC113320 Mus muscu
c 959 38 79.2 267904 14 AC156870 AC156870 Bos tauru
c 960 38 79.2 270337 14 AC112074 AC112074 Rattus no
c 961 38 79.2 270378 14 AC096495 AC096495 Rattus no
c 962 38 79.2 270835 14 AC160643 AC160643 Gallus ga
c 963 38 79.2 271368 14 AC134642 AC134642 Rattus no
c 964 38 79.2 273527 14 AC156740 AC156740 Bos tauru
c 965 38 79.2 273606 14 AC116272 AC116272 Rattus no
c 966 38 79.2 274026 14 AC129633 AC129633 Rattus no
c 967 38 79.2 275923 14 AC154571 AC154571 Mus muscu
c 968 38 79.2 276624 14 AC132560 AC132560 Rattus no
c 969 38 79.2 281574 14 AC023604 AC023604 Mus muscu
c 970 38 79.2 283507 14 AC152628 AC152628 Bos tauru
c 971 38 79.2 283507 14 AC152628 AC152628 Bos tauru
c 972 38 79.2 287055 14 AC155281 AC155281 Bos tauru
c 973 38 79.2 287328 14 AC112394 AC112394 Rattus no
c 974 38 79.2 288358 14 AC106121 AC106121 Rattus no
c 975 38 79.2 289107 14 AC152440 AC152440 Bos tauru
c 976 38 79.2 294706 14 AC157182 AC157182 Bos tauru
c 977 38 79.2 295150 14 AC0939125 AC0939125 Streptomy
c 978 38 79.2 295400 14 AC103344 AC103344 Rattus no
c 979 38 79.2 295584 14 AC117896 AC117896 Rattus no
c 980 38 79.2 300703 14 AC015940 AC015940 Clostridi
c 981 38 79.2 301337 15 AC017082 AC017082 Oryza sat
c 982 38 79.2 302506 14 AC111719 AC111719 Rattus no
c 983 38 79.2 316900 14 AC111719 AC111719 Rattus no
c 984 38 79.2 341944 14 AC079158 AC079158 Homo sapi
c 985 38 79.2 349746 14 BX572099 BX572099 Prochloro
c 986 37 77.1 119 1 LARNSS X05041 Listonella
c 987 37 77.1 120 1 LTRNRS X06688 Listonella
c 988 37 77.1 120 1 VMRNRS X15179 Vibrio medi
c 989 37 77.1 120 6 AR427924 AR427924 Sequence
c 990 37 77.1 122 1 BHRNRS X02030 Beneckea ha
c 991 37 77.1 240 15 AF024109 AF024109 Oryza sat
c 992 37 77.1 289 8 HS7F6F Z63238 H. sapiens C
c 993 37 77.1 328 6 AX246821 AX246821 Sequence
c 994 37 77.1 347 2 AF320428 AF320428 Myrsidea
c 995 37 77.1 438 3 AF010079 AF010079 Unidentif
c 996 37 77.1 445 6 AR240393 AR240393 Sequence
c 997 37 77.1 456 10 BV290748 BV290748 S232P6152
c 998 37 77.1 470 10 BV222125 BV222125 S233P641R
c 999 37 77.1 509 10 BV293188 BV293188 S232P676R
c1000 37 77.1 531 3 AB075785 AB075785 Unculture

```

ALIGNMENTS

```

RESULT 1
CQ920916 475 bp DNA linear PAT 23-NOV-2004
LOCUS
DEFINITION Sequence 2116 from Patent WO2004097052.
ACCESSION CQ920916
VERSION CQ920916.1 GI:56210857
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Burczynski, M.E., Twine, N.C., Slonim, D.K., Trepicchio, W.L.,
Strahs, A., Immerman, F. and Dörner, A.J.
TITLE Methods for prognosis and treatment of solid tumors
JOURNAL Patent: WO 2004097052-A 2116 11-NOV-2004;
Wysht (US); Burczynski, Michael E. (US)
FEATURES
Location/Qualifiers
source
1..475
/organism="Homo sapiens"

```

```

/mot_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 2.42 Length: 475
Pred. No.: 48.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-9 (1-9) x CQ920916 (1-475)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 126 GGTCTACCCACATTAGGTTTCCTG 152

RESULT 2
AX829164 927 bp DNA linear PAT 12-DEC-2003
LOCUS
DEFINITION Sequence 57 from Patent WO2059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Mack, D.H., Gish, K.C. and Afar, D.
TITLE Methods of diagnosis of breast cancer, compositions and methods of
screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)
FEATURES
Location/Qualifiers
source
1..927
/organism="Homo sapiens"
/mot_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 4.86 Length: 927
Pred. No.: 48.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-9 (1-9) x AX829164 (1-927)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 505 GGTCTACCCACATTAGGTTTCCTG 531

RESULT 3
BD249731 1263 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1263)
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;

```

OXFORD BIOMEDICA LTD
 OS Homo sapiens (human)
 PN JP 2002530060-A/1
 PD 17-SEP-2002
 PF 18-NOV-1999 JP 200582415
 PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
 30-JUL-1999 GB 9917995.4
 PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
 PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K7/06, C07K14/065,
 C07K19/00,
 PC C12N15/00
 CC Polypeptide
 FH Key Location/Qualifiers
 FT source 1..1263 /organism="Homo sapiens (human)"
 FT
 FT
 FT
 FEATURES source
 1..1263
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.71 Length: 1263
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-9 (1-9) x BD249731 (1-1263)
 QY 1 GlyLeuProHisIleArgValPheLeu 9
 |||||
 DB 847 GGTCTACCCACCATAGGGTTTCCTG 873
 RESULT 4
 LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
 DEFINITION Sequence 1 from Patent WO029428.
 ACCESSION AX025011
 VERSION AX025011.1 GI:10184932
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Carroll, M.W. and Myers, K.A.
 TITLE Polypeptide
 JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
 CARROLL, MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
 BIOMEDICA LTD (GB)
 FEATURES source
 1..1263
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.71 Length: 1263
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-9 (1-9) x AX025011 (1-1263)
 QY 1 GlyLeuProHisIleArgValPheLeu 9
 |||||
 DB 847 GGTCTACCCACCATAGGGTTTCCTG 873

DB 847 GGTCTACCCACCATAGGGTTTCCTG 873
 RESULT 5
 LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
 DEFINITION Sequence 1 from Patent EP1160323.
 ACCESSION AX316086
 VERSION AX316086.1 GI:17899278
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Carroll, M.W. and Myers, K.A.
 TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
 JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
 Oxford Biomedica (UK) Limited (GB)
 FEATURES source
 1..1263
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.71 Length: 1263
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-9 (1-9) x AX316086 (1-1263)
 QY 1 GlyLeuProHisIleArgValPheLeu 9
 |||||
 DB 847 GGTCTACCCACCATAGGGTTTCCTG 873
 RESULT 6
 LOCUS CQ731678 2053 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 17612 from Patent WO2068579.
 ACCESSION CQ731678
 VERSION CQ731678.1 GI:42308932
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 human exons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 17612 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES source
 1..2053
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 11.1 Length: 2053
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

```

US-10-774-176-9 (1-9) x CQ731678 (1-2053)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 931 GGTCTACCCACATTAGGGTTTCCTG 957

RESULT 7
HS5T4OA 2053 bp RNA linear PRI 18-APR-2005
LOCUS Homo sapiens 5T4 gene for 5T4 oncofetal antigen.
ACCESSION Z29083
VERSION Z29083.1 GI:435654
KEYWORDS 5T4 gene; 5T4 oncofetal antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2053)
AUTHORS Myers,K.A., Rahi-Saund,V., Davison,M.D., Young,J.A., Chester,A.J.
and Stern,P.L.
TITLE Isolation of a cDNA encoding 5T4 oncofetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
J. Biol. Chem. 269 (12), 9319-9324 (1994)
JOURNAL PUBMED 8132670
REFERENCE 2 (bases 1 to 2053)
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK

FEATURES
source
1..2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label=N-flank
85..1347
/codon_start=1
/evidence=experimental
/product="5T4 oncofetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCSGPAAGDGLRLRLARLALLVGLWSSSTSSASSPSS
APFLASVASQPLPDPQPCALCESEAAATKVCNRLNTEVPTDLPAYRNLFLTGNQ
LAVLPAGARPPPLALNALNLGSLRDLDEVRAGAFHLPSLRQLDLSHNPLADLSPF
AFSGNSASVSPSLVLEILNHI VPPDERONRSPFGMVVAALLAGRALQGLRLLELA
SNHYLPDVLQIPLSLRLDLNNLSVLSLYSPRNLTLSLHLEEDNALKVLNG
TLASLQGLPHLRVFDLNNPWCDCMADMTLKEIVVQKDRLTCAYPEKRNRLV
LELNSADLDCDQPIPLPSPSLQTSYFVIGLVALIGALFLDLVLYLNRRGIKKWHNINDAC
RDHMEGHYRYEINADPRLTNLSNSDV"
130..171
/sig_peptide
misc_RNA 373..966
/product="Leucine rich repeat region"
/label=LRRS
966..1119
/misc_RNA 966..1119
/product="LRR C-terminal flank"
/label=C-flank
1153..1215
/misc_RNA 1153..1215
/product="transmembrane peptide"
/standard_name="transmembrane region"

/function="Anchorage of the protein to the cell membrane"

ORIGIN
Alignment Scores: 11.1 Length: 2053
Pred. No.: 48.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 8

US-10-774-176-9 (1-9) x HS5T4OA (1-2053)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 931 GGTCTACCCACATTAGGGTTTCCTG 957

RESULT 8
BD127282 2359 bp DNA linear PAT 18-SEP-2002
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD127282
ACCESSION BD127282
VERSION BD127282.1 GI:23222227
KEYWORDS JP 2002017375-A/2713
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Oeuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2713 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2713
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers
1..2359
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1..2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 12.9 Length: 2359
Pred. No.: 48.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-9 (1-9) x BD127282 (1-2359)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 1270 GGTCTACCCACATTAGGGTTTCCTG 1296

```

```

RESULT 9
CQ782724          CQ782724          2359 bp      DNA      linear      PAT 17-MAR-2004
DEFINITION       Sequence 2864 from Patent EP1396543.
ACCESSION        CQ782724
KEYWORDS         CQ782724.1  GI:45502667
SOURCE           Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS          Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE            Primers for synthesizing full length cDNA clones and their use
JOURNAL          Patent: EP 1396543-A 2864 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES         Location/Qualifiers
source           1..2359
                 /organism="Homo sapiens"
                 /mol_type="unassigned DNA"
                 /db_xref="taxon:9606"
                 424..1575
                 /note="unnamed protein product"
                 /codon_start=1
                 /protein_id="CAF85958.1"
                 /db_xref="GI:45502668"
                 /translation="MPGCSRGPAAGDGLRLARLALVLGWSSSPTSSASSPSSS
APPLASVASOPPLDQCPALCESEARTVKVNRNLTETPTDLPAYVRNLFITGNQ
LAVLPAGAPRPPLEALALNLSGRLDDEVAGAFELPSLRQLDLSHNPADIUSPF
AFSGNSAVSAPSLVELILNHI VPEDEQRNSPEGMVAALAGRALQGLRLLELA
SNHFLYLPDRYLAQPLSHLDLNNLSVLTYSFRNLTHLESJLHEDNALKVLHNG
TLAEIQLGPHIRVFLDNNPVCDCMDMVTWLTETVVGKDRLTCAYPEKMRNRVL
LELSADLDCDPLPPSLQTSVFIGLIVLIGALFLLVLYLNRKGKK"
CDS
ORIGIN
Alignment Scores:
Pred. No.:      12.9      Length:      2359
Score:          48.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             6        Gaps:        0

US-10-774-176-9 (1-9) x CQ782724 (1-2359)

Qy      1  GlyLeuProHisIleArgValPheLeu 9
Db      1270  GGTCTACCCACATTAGGGTTTTCCTG 1296

RESULT 10
AK074786          AK074786          2359 bp      mRNA      linear      PRI 03-SEP-2002
DEFINITION       Homo sapiens cDNA FLJ90305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION        AK074786
KEYWORDS         AK074786.1  GI:22760460
SOURCE           Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS          Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
TITLE            NEDO human cDNA sequencing project

```

```

Unpublished
2 (bases 1 to 2359)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES         Location/Qualifiers
source           1..2359
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="NT2RP2000694"
                 /cell_line="NT2"
                 /cell_type="teratocarcinoma"
                 /clone_lib="NT2RP2"
                 /note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"
ORIGIN
Alignment Scores:
Pred. No.:      12.9      Length:      2359
Score:          48.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             8        Gaps:        0

US-10-774-176-9 (1-9) x AK074786 (1-2359)

Qy      1  GlyLeuProHisIleArgValPheLeu 9
Db      1270  GGTCTACCCACATTAGGGTTTTCCTG 1296

RESULT 11
BD127283          BD127283          2361 bp      DNA      linear      PAT 18-SEP-2002
LOCUS            BD127283
DEFINITION       Primer for synthesizing full-length cDNA and use thereof.
ACCESSION        BD127283
VERSION          BD127283.1  GI:23222228
KEYWORDS         JP 2002017375-A/2714.
SOURCE           Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS          Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE            Primer for synthesizing full-length cDNA and use thereof
JOURNAL          Patent: JP 2002017375-A 2714 22-JAN-2002;
HELIX RESEARCH INSTITUTE
COMMENT          OS Homo sapiens (human)
PN JP 2002017375-A/2714
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,

```

PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
(426)...(1685).

FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 12.9 Length: 2361
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-9 (1-9) x BD127283 (1-2361)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
|||||
Db 1272 GGTCTACCCACATTAGGGTTTCCTG 1298

RESULT 12
CQ782726
LOCUS CQ782726 2361 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2866 from Patent EP1396543.
ACCESSION CQ782726
VERSION CQ782726.1 GI:45502669

KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.

TITLE
JOURNAL Primers for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 2866 10-MAR-2004;
Research Association for Biotechnology (JP)

FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS
426..1688
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85961.1"
/db_xref="GI:45502670"
/translation="MPGCSRGPAAGDGLRLRLALVLLGWSSSPSSASSPSS
APFLASVASQPPPLDQPCALCESEAAITVKCNRNLTVEPTDLPAVYRNLFITGNQ
LAVLPAGAFARPLAEALNLSGRDLDEVRAGAFELPSLRQLDLSHNLADLSPP
AFSGNSASVAPSLVELILNHI VPPDERQNRSPGVMVAALLAGRALQGLRLLELA
SNHFLYPRDVLALQPSLRHLDLNNLSVSLTYVSFRNLTHLESILHLEDNALKVLHNG
TLAEGLPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPEKMRNVL
LELMSADLDCDFILPSPSQTSYVFGIVLALIGAIFLLVLYLNRRGIKKMMNIRDAC
RDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 12.9 Length: 2361
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-9 (1-9) x CQ782726 (1-2361)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
|||||
Db 1272 GGTCTACCCACATTAGGGTTTCCTG 1298

RESULT 13
AX961916

LOCUS AX961916 2361 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 127 from Patent WO03104277.
ACCESSION AX961916
VERSION AX961916.1 GI:40881326

KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.

AUTHORS
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 127 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)

FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS
426..1688
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06467.1"
/db_xref="GI:40881327"

/translation="MPGCSRGPAAGDGLRLRLALVLLGWSSSPSSASSPSS
APFLASVASQPPPLDQPCALCESEAAITVKCNRNLTVEPTDLPAVYRNLFITGNQ
LAVLPAGAFARPLAEALNLSGRDLDEVRAGAFELPSLRQLDLSHNLADLSPP
AFSGNSASVAPSLVELILNHI VPPDERQNRSPGVMVAALLAGRALQGLRLLELA
SNHFLYPRDVLALQPSLRHLDLNNLSVSLTYVSFRNLTHLESILHLEDNALKVLHNG
TLAEGLPHIRVFLDNNPWCDCMADMTWLKETEYVQGDRLTCAYPEKMRNVL
LELMSADLDCDFILPSPSQTSYVFGIVLALIGAIFLLVLYLNRRGIKKMMNIRDAC
RDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 12.9 Length: 2361
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-9 (1-9) x AX961916 (1-2361)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
|||||
Db 1272 GGTCTACCCACATTAGGGTTTCCTG 1298

RESULT 14
AK074790

LOCUS AK074790 2361 bp mRNA linear PRI 09-JUL-2005
DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.

ACCESSION AK074790

VERSION AK074790.1 GI:22760466

KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

1 Otsuki, T., Ota, T., Nishikawa, T., Hayashi, K., Suzuki, Y.,
Yamamoto, J., Wakamatsu, A., Kimura, K., Sakamoto, K., Hatano, N.,
Kawai, Y., Ishii, S., Saito, K., Kojima, S., Sugiyama, T., Ota, T.,

Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Nagai, K., Sugano, S. and Isoqai, T.
Signal Sequence and Keyword Trap in silico for Selection of Full-length Human cDNAs Encoding Secretion or Membrane Proteins from Oligo-Capped cDNA Libraries
DNA Res. 12, 117-126 (2005)

2
Isoqai, T., Ota, T., Nishikawa, T., Hayaashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagahara, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Ninomiya, K.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2361)
Isoqai, T. and Otsuki, T.
Direct Submission
Submitted (25-MAR-2002) Takao Isoqai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- and 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).

TITLE
NEDO human cDNA sequencing project

JOURNAL
Unpublished

REFERENCE
3 (bases 1 to 2361)

AUTHORS
Isoqai, T. and Otsuki, T.

TITLE
Direct Submission

JOURNAL
Submitted (25-MAR-2002) Takao Isoqai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- and 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).

FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 12.9 Length: 2361
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x AK074790 (1-2361)

QY 1 GlyLeuProHisIleArgValPheLeu 9
Db 1272 GGTCTACCCACCATGAGGTTCCTG 1298
|||||
BC037161 2379 bp mRNA linear PRI 29-JUN-2004
LOCUS BC037161
DEFINITION Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317 IMAGE:4138906), complete cds.
ACCESSION BC037161
VERSION BC037161.2 GI:33872201
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2379)
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F., Diatchenko, L., Marusina, K., Parner, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, I.L., Carninci, P., Prange, C., Raha, S., Usdin, T.B., Toshiyuki, S., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 2379)
Strausberg, R.
Direct Submission
Submitted (03-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:22713382.
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Plate: 26 Row: m Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5729717.

FEATURES
source
1..2379
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:15317 IMAGE:4138906"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2379
/gene="TPBG"
/note="synonyms: MGP1, 5T4-AG, 5T4"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"
427..1689
/gene="TPBG"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="AAH37161.1"
/db_xref="GI:22713383"

CDS

```

/db_xref="GeneID:7162"
/db_xref="MIM:190920"
/translation="MPGCSRGPAAGDQGLRLARLALVLLGWSSSPTSSASSPSS
APFLASVSAQPPPLDQPCALCESEARTVKCVNRNLTEVPTDLPAYVRNLFLTGNQ
LAVLPAGAFARPPPLAEALNLSGSRLEDEVAGAFEHLPRLQDLDSHNPLADLSPF
AFSGNSASVSPPLVELILNHI VPPEDERQNRSPGVMVAALLAGRALQGLRLSLA
SNHFLYLPDVLQALPSRLHLDLNNLSVLSLVTSFRNLTHLESLEHLEDAKLVHNG
TLAEQLQGLPHIRVFLDNNPWCDCMADMTWLKETEVOQKDRLTCAYPEKMRNVL
LELNSADLDCDPLPSPQTSYVFGIVLALIGAFLLVLYLNRRGIKKWMHNRDAD
RDHMEGYHYRYEINADPRLTNLSNSDV"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 13 Length: 2379
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

```

US-10-774-176-9 (1-9) x BC037161 (1-2379)

```

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 1273 GGTCTACCCACATTAGGCTTTTCCTG 1299

```

RESULT 16

```

HSA012159 HSA012159 5551 bp DNA linear PRI 15-APR-2005
LOCUS HOMO sapiens 5T4 oncofetal trophoblast glycoprotein gene.
DEFINITION
ACCESSION AJ012159
VERSION AJ012159.1 GI:3805946
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE HOMO sapiens
ORGANISM HOMO sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

```

REFERENCE

```

AUTHORS King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10366710
REFERENCE 2 (bases 1 to 5551)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK

```

FEATURES

```

source
1..5551
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
2698..2703
/bound_moiety="Sp1"
2704..2709
/bound_moiety="Sp1"
2716..27400
/gene="5T4"
2716..2800
/gene="5T4"
/evidence=experimental
2801..3092
/gene="5T4"
/evidence=experimental
3093..5400
/gene="5T4"
/evidence=experimental
3431..4693
/gene="5T4"

```

```

/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA0930.1"
/db_xref="GI:3805947"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCSRGPAAGDQGLRLARLALVLLGWSSSPTSSASSPSS
APFLASVSAQPPPLDQPCALCESEARTVKCVNRNLTEVPTDLPAYVRNLFLTGNQ
LAVLPAGAFARPPPLAEALNLSGSRLEDEVAGAFEHLPRLQDLDSHNPLADLSPF
AFSGNSASVSPPLVELILNHI VPPEDERQNRSPGVMVAALLAGRALQGLRLSLA
SNHFLYLPDVLQALPSRLHLDLNNLSVLSLVTSFRNLTHLESLEHLEDAKLVHNG
TLAEQLQGLPHIRVFLDNNPWCDCMADMTWLKETEVOQKDRLTCAYPEKMRNVL
LELNSADLDCDPLPSPQTSYVFGIVLALIGAFLLVLYLNRRGIKKWMHNRDAD
RDHMEGYHYRYEINADPRLTNLSNSDV"
sig_peptide 3431..3516
mat_peptide 3517..4690
polyA_signal 5331..5336
polyA_signal 5380..5385
ORIGIN
Alignment Scores:
Pred. No.: 31.3 Length: 5551
Score: 48.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-10-774-176-9 (1-9) x HSA012159 (1-5551)

```

Qy 1 GlyLeuProHisIleArgValPheLeu 9

```

Db 4277 GGTCTACCCACATTAGGCTTTTCCTG 4303

```

RESULT 17

```

HSA0492P14 HSA0492P14 121909 bp DNA linear PRI 18-MAY-2005
LOCUS HUMAN DNA sequence from clone RP3-492P14 on chromosome 6q13-15
DEFINITION Contains a single stranded DNA binding protein pseudogene, the TPBG
gene for trophoblast glycoprotein (5T4-AG) and a CpG island,
complete sequence.
ACCESSION AL121977
VERSION AL121977.11 GI:11863678
KEYWORDS HTG; CpG island; TPBG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 121909)
AUTHORS Garner, P.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Dec 15, 2000 this sequence version replaced gi:11558491.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Swi, SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at

```


<http://www.sanger.ac.uk/HGP/Chr6>
 RP3-492P14 is from the library RPCI-3 constructed by the group of
 Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pCYPAC2
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: vega@sanger.ac.uk

 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one subclone; and the assembly was confirmed by restriction digest,
 except on the rare occasion of the clone being a YAC.

FEATURES

SOURCE

1. 121909
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="q13-15"
 /clone="RP3-492P14"
 /clone_lib="RPCI-3"
 100

misc_feature

/note="Clone right end: RP1-93K22"
 complement(10004..10982)
 /locus_tag="RP3-492P14.2-001"
 /pseudoc

CDS

complement(10004..10982)
 /locus_tag="RP3-492P14.2-001"
 /note="match: proteins: F81877 Q99LX9 Q9BWW6 Q9CVZ8 Q9D6L4
 Q9P038 Q9Y4T7"
 /pseudoc

misc_feature

/codon_start=1
 86539
 /note="Clone left end: RP1-90G1"
 109639..116836
 /gene="TPBG"

gene

Join(109639..109916,110631..116836)
 /gene="TPBG"

mRNA

Join(109639..109916,110631..116836)
 /locus_tag="RP3-492P14.1-001"
 /note="match: ESTs: AA149121 AA152323 AA565852 AA643734
 AL544610 AW471072 AW62538 BE260089 BF306457 BF306926
 BF314984 B1196133 B1562387 BM069633 BM670613
 match: cDNAs: AJ420536.1 Z29083.1"
 110970..112232
 /gene="TPBG"

CDS

/locus_tag="RP3-492P14.1-001"
 /standard_names="OTTHUMP0000016786"
 /note="match: proteins: Q13641 Q9QYD9 Q9ZOL0"
 /codon_start=1
 /product="trophoblast glycoprotein"

/protein_id="CAI21546.1"
 /db_xref="GI:56203539"
 /db_xref="Genew:12004"
 /db_xref="GOA:Q13641"
 /db_xref="InterPro:IPR000172"
 /db_xref="InterPro:IPR000483"
 /db_xref="InterPro:IPR001611"
 /db_xref="InterPro:IPR003591"
 /db_xref="UniProt/TreMBL:Q13641"
 /translation="MPGCSRGPAAGDGRRLRLALVLGVSSSSPTSSASSFSSS
 APFLASVAPPLPDQPCALCESEARTKVNRLTEVPTDLPAVYRNPLFGNQ
 LAVLPAGAPARRPFLAELNALNSGRLEDRVAGAPEHLPSLRQLDLSHNPADI.SPF
 APGSGNASVSPPLVEILNHHVPPDERQNRSPFGMVVAALLAGRALQGLRRLELA
 SNHFLYPRDLVAQLPSLRHLDLSNLSVTSVFRNLTHLESRLHLEDNALKVLHNG
 TLAEIQGLPHIRVFLDNNPWCDCIMADMTWLKETEVVQGRDLTCAYPEKVRNRL

LELNSADLDCDILPPSLQTSVFLGIVLALGAIFFLLVLYLNKRGKKMWHNRDAC
 RDHMGYHYRYEINADPRITLNLSSNDV"
 116817..116822
 /gene="TPBG"
 /locus_tag="RP3-492P14.1-001"
 116836
 /gene="TPBG"
 /locus_tag="RP3-492P14.1-001"
 121909
 /note="Clone right end: RP3-492P14"

ORIGIN

Alignment Scores:
 Pred. No.: 782 Length: 121909
 Score: 48.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-9 (1-9) x HSJ492P14 (1-121909)

QY 1 GlyLeuProHisIleArgValPheLeu 9
 |||
 DB 111816 GGTCTACCCACATAGGGTTTCCTG 111842

RESULT 18

AB168308
 LOCUS
 DEFINITION
 Macaca fascicularis testis cDNA clone: QsEA-11109, similar to human
 trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.
 AB168308
 ACCESSION
 VERSION
 AB168308.1 GI:67967899
 KEYWORDS
 SOURCE
 Macaca fascicularis (crab-eating macaque)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE

1
 AUTHORS
 TITLE
 DNA sequences of macaque genes expressed in brain of testis and its
 evolutionary implications
 JOURNAL
 REFERENCE
 Unpublished
 AUTHORS

TITILE

Osada, N., Hirata, M., Tanuma, R., Kusuda, J., Hida, M., Suzuki, Y.,
 Sugano, S., Gojobori, T., Shen, J. C.-K., Wu, C. I. and Hashimoto, K.
 Substitution rate and structural divergence of 5'UTR evolution:
 Comparative analysis between human and cynomolgus monkey cDNAs
 Unpublished
 3 (bases 1 to 2714)
 Hashimoto, K., Kusuda, J. and Sugano, S.
 Direct Submission

JOURNAL

Submitted (18-MAR-2004) Katsuyuki Hashimoto, National Institute of
 Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
 1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan
 (E-mail: khashi@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genebank/>,
 Tel: 81-3-5285-1111 (ex.2120), Fax: 81-3-5285-1181)
 The international consortium for macaque cDNA sequencing and
 analysis consists of: Department of Virology and Human Genome
 Center, Institute of Medical Science, The University of Tokyo,
 Tokyo, Japan; Division of Genetic Resources, National Institute of
 Infectious Diseases of Japan, Tokyo, Japan; National Health
 Research Institute, Taipei, Taiwan; Institute of Molecular Biology,
 Academia Sinica, Taipei, Taiwan; Department of Ecology & Evolution,
 University of Chicago, Chicago, IL, USA; Center for Information
 Biology, National Institute of Genetics of Japan, Mishima, Japan.
 Clone distribution: clone distribution information can be found at:
<http://www.nih.go.jp/yoken/genebank/>

COMMENT

Lab host: TOP10
 Vector: PMB18S-FL3 (Acc.No. AB009864)
 R. Site1: DraIII (CACTGTGTG)
 R. Site2: DraIII (CACCATGTG)

troglydites verus), 3 other Pan troglodytes verus chimps (Donald,Karlien,Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin (Gon,Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNQS(30,25)(single strand NQS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNQS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NQS(30,25) standard was applied to all pairs of overlapping reads to call NQS bases and SNPs. Alignments (between two reads) with less than 100 NQS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

FEATURES

source
1..779
/organism="Pan troglodytes troglodytes"
/mol_type="genomic DNA"
/sub_species="troglodytes"
/db_xref="taxon:37011"
/clone_lib="Clara"
<1..>779

STS

Alignment Scores:
Pred. No.: 45.3 Length: 779
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.6% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-9 (1-9) x BV595106 (1-779)

Cy 1 GlyLeuProHisIleArgValPheLeu 9
Db 263 GGCCTCCACATCTGAAGTTTITG 237

RESULT 24
AF014939 LOCUS 24950 bp DNA linear INV 27-MAY-2005
DEFINITION Caenorhabditis elegans cosmid ZC132, complete sequence.
ACCESSION AF014939
VERSION AF014939.1 GI:2275620
KEYWORDS HTG.

SOURCE
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 24950)

REFERENCE
AUTHORS
CONSTRM
TITLE

C. elegans Sequencing Consortium
Genome sequence of the nematode C. elegans: a platform for
investigating biology
JOURNAL Science 282 (5396), 2012-2018 (1998)
PUBMED 9851916

REFERENCE
AUTHORS
Bradshaw,H. and Devlin,K.

TITLE
JOURNAL
REFERENCE
AUTHORS

The sequence of C. elegans cosmid ZC132
Unpublished (2001)
3 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (18-JUL-1997) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

4 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (29-MAY-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

5 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (22-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

6 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (22-SEP-2004) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

7 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (19-JAN-2005) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

8 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (12-FEB-2005) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

9 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (19-MAY-2005) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

10 (bases 1 to 24950)
Waterston,R.
Direct Submission
Submitted (27-MAY-2005) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

REFERENCE
AUTHORS

Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: submissions@watson.wustl.edu and jess@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate chemistry
or covered by high quality data (i.e., phased quality >= 30); an
attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by sequence from

more than one mi3 subclone.

For a graphical representation of this clone sequence and its analysis see:
http://www.wormbase.org/db/seq/sequence?name=ZC132;class=Sequence

NEIGHBORING CLONE INFORMATION

The 5' clone is 847D2, 200 bp overlap; the 3' clone is R02D1, 950 bp overlap. Actual start of this clone is at base position 29759 of 847D2; actual end is at 24950 of ZC132.

NOTES:

Coding sequences below are the result of integration and manual review of the following data : computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CS_INDEX.html) and The C. elegans ORFome cloning project (http://wormdb.dfc.harvard.edu/), similarity to other proteins from BlastX analyses (http://blast.wustl.edu/), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. TRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES
source

Location/Qualifiers
1. 24950
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="v"
/clone="ZC132"
complement(326..3370)
/locus_tag="ZC132.3"

gene

CDS
complement(join(326..451,499..693,1001..1399,1448..1633,2232..2276,2323..2517,2729..3127,3179..3370))
/locus_tag="ZC132.3"
/standard_name="ZC132.3a"
/note="contains similarity to Staphylococcus epidermidis Cytochrome d ubiquinol oxidase subunit II-like protein.; TR:Q8CPN5; coded for by the following C. elegans cDNAs: OSTF188A4.1, OSTF188A4.1, YK783a01.5, YK783a01.3"
/codon_start=1
/product="Hypothetical protein ZC132.3a"
/protein_id="AAB63926.1"
/db_xref="GI:2275623"

/db_xref="WormBase:ZC132.3a"
/translation="MVRKIKLSIFPYSLSDVKFVDFTIGNVSHVFLGIFLGLVVFYGLIIHMFQKSECEQDQSLKPLLSIYRTITKRIFTVHGHTLVNGLSHFNVAGLVTVMVPMWALYSIKVPTMYLIVISPPSISRCFVPSLAKPSVELTQCVKAGIRFVLGVLKDLVLSLIVSEAKPEKIGIIDIYCGIYLFQCLLLFVAGLSQLPIMREKSAEIPHEKLIYSHTPAAILKLVLLISLVSGSFNIEIIMSIFVSDSLFVPTQMKILNIPYSLSFDRNITKRIIVTVNGIALVNGLSHFVGMVLTVPWMTALYSFVQDSFTPLISKRNITKRIIVTVNGIALVNGLSHFVGMVLTVPWMTALYSIKVPTMYLIVISPPSISRCFVPSLAKPSVELTQCVKAGIRFVLGVLKDLVLYSLIIVSEAKPEKIGIIDIYCGIYLFQCLFLCVAGLSQLPIMCKEISIELTHEKQIYSQTPAAILKLVLLFCLFVAGNFNIEIIMSIFVLSLFPVSMVLIFVFCRPKHVEQVEMVTHQA"
complement(4846..7947)
/locus_tag="ZC132.4"

gene

CDS
complement(join(4846..5712,5882..6214,6652..7947))
/locus_tag="ZC132.4"
/standard_name="ZC132.4a"
/note="contains similarity to Ukuniemi virus M polyprotein precursor [Contains: Glycoprotein GI; Glycoprotein G2]; SW:P09613"
/codon_start=1
/product="Hypothetical protein ZC132.4a"
/protein_id="AAB63927.1"
/db_xref="GI:2275624"

/db_xref="WormBase:ZC132.4"
/translation="MMIHLFRVPVPIGAPISMAWSVIKRTISTIWNLGRACLRTRR AQRHSWSTIIPILCIIVTVSSGCOEISLQSOIKICSTDDGTSYCKLVTRPMANIRP YKSCIRITKSGKVTKEVRQKAVOLECMKTTVTFTKDVQVTKSSKRCRCHPMSCG EDKNNISLSLPLEBTNGVETKRCSESCGCGCGCFYSSGCLFRIIYAVPTIN DSFEIMKNEIWHKLVKSTEDMGKPRMSEQIKIPLNFVTIICDITMSVNSLTK CQTESLQNMNSLDKRLPIQENHMKLQSLDDKITATIANEVTVTMTLKDNVTVSAW ISNEKYVKAASHAGCYKATGSSATLQCFQKTDITIGNIYVWKSVQVIRNKRRLRRS TRAFKKQKPCIKTKNKHQLFSETRKNLNNPAAQAVDQDADHDIIELELPQAP TSQNGDDKIDQEHSDRSQAHKDLQKQGNKRNKLAIVTASDMMLPHHQLNWA QKDPARSRLKQKQRFNVHQALNTLSANKDGTSSKMSLSDMKKEAWTQIQAHMT QSPENRLEHQIQILQKQDQEMRSIRLILETSLKRDILNHPKQNEAISMEVINHEE SERLPRTDLQFNLRNYKIPKIKBPTLQKSGSQPLKAGENLSQIKPIKQGNKML FKAFHRSPTVADREPSPSCAVCTRHLSDRGIIKDKPKARKAHLENNRCLLCLKK BELHBFQKRCCHYCHRTNHSWCLQAPISGVH"
complement(8181..12250)
/locus_tag="ZC132.5"
complement(join(8181..9060,9297..12250))
/locus_tag="ZC132.5"
/standard_name="ZC132.5a"
/note="contains similarity to Pfam domains PF00665 (integrase core domain), PF05380 (Pao retrotransposon peptidase)"
/codon_start=1
/product="Hypothetical protein ZC132.5"
/protein_id="AAB63932.1"
/db_xref="GI:2275629"

gene

CDS
complement(join(8181..9060,9297..12250))
/locus_tag="ZC132.5"

/standard_name="ZC132.5"
/note="contains similarity to Pfam domains PF00665 (integrase core domain), PF05380 (Pao retrotransposon peptidase)"
/codon_start=1
/product="Hypothetical protein ZC132.5"
/protein_id="AAB63932.1"
/db_xref="GI:2275629"
/db_xref="WormBase:ZC132.5"
/translation="MTNEDLAFALTKOLKINOECFETRVDPDIIGCOHIAIEIRGDF IQLSPGILLIKTVFGYTMGNTKWSKPLEMDSTISVMTVVKVQNDDEIFLQOQETI MHAPNEITGAVVDEKLEMEKITTQFPNTIKRENGYHVLFPKKEVIDKLPSEFALIA KRLQSGUKANPQVKLVNDVFQISKNILREVDVSKDTGMRRIHYNPHSPVLIPQK TTKCRVVIDGSAFKPNPSLNDALYQFTILPDSVDFRSGKTVLLADVSKAFQV HLNESDRDVTVLVAVNPDLPTRENLEVLFTVLFGVSPFLGATILFHLDRME DKCLANTARNLYVDNLIIATDDSEAMFKLYNKVTVFNGLSMNIREFQSNDSQFTD

Alignment Scores:

Pred. No.: 1.67e+03 Length: 24950
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 89.6% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-9 (1-9) x AF014939 (1-24950)

QY 1 GlyLeuProHisIleArgValPhe 8

Db 3181 GGTCTCTCTCATTCGGATTTT 3204

RESULT 25

AC025304/c 72957 bp DNA linear HTG 13-JUL-2000
AC025304
AC025304
AC025304.1 GI:7209941
HTG; HTGS PHASE0.
Homo sapiens (human)
Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 72957)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-23P8
Unpublished
2 (bases 1 to 72957)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,

REFERENCE

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivari, T. M., Oliver, J., Peterson, K., Pierre, N., Plesani, C., Pollara, V., Raymond, C., Riley, P., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggilio, J., Vassilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

All repeats were identified using RepeatMasker:

Smit, A.P.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4416

Center clone name: 23_P_8

* NOTE: This record contains 88 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 * 724: contig of 724 bp in length
 * 725
 * 824: gap of 100 bp
 * 825
 * 1557: contig of 733 bp in length
 * 1558
 * 1657: gap of 100 bp
 * 1658
 * 2391: contig of 734 bp in length
 * 2392
 * 2491: gap of 100 bp
 * 2492
 * 3229: contig of 738 bp in length
 * 3230
 * 3329: gap of 100 bp
 * 3330
 * 4050: contig of 721 bp in length
 * 4051
 * 4150: gap of 100 bp
 * 4151
 * 4882: contig of 732 bp in length
 * 4883
 * 4982: gap of 100 bp
 * 4983
 * 5698: contig of 716 bp in length
 * 5699
 * 5798: gap of 100 bp
 * 5799
 * 6525: contig of 727 bp in length
 * 6526
 * 6625: gap of 100 bp
 * 6626
 * 7354: contig of 729 bp in length
 * 7355
 * 7454: gap of 100 bp
 * 7455
 * 8189: contig of 735 bp in length
 * 8190
 * 8289: gap of 100 bp
 * 8290
 * 9035: contig of 746 bp in length
 * 9036
 * 9135: gap of 100 bp
 * 9136
 * 9885: contig of 750 bp in length
 * 9886
 * 9885: gap of 100 bp
 * 9986
 * 10733: contig of 748 bp in length
 * 10734
 * 10833: gap of 100 bp
 * 10834
 * 11555: contig of 722 bp in length
 * 11556
 * 11655: gap of 100 bp
 * 11656
 * 12378: contig of 723 bp in length
 * 12379
 * 12478: gap of 100 bp
 * 12479
 * 13209: contig of 731 bp in length

13210
 * 13309: gap of 100 bp
 * 13310
 * 14023: contig of 714 bp in length
 * 14024
 * 14123: gap of 100 bp
 * 14124
 * 14669: contig of 746 bp in length
 * 14670
 * 14969: gap of 100 bp
 * 15715
 * 15715: contig of 746 bp in length
 * 15716
 * 15815: gap of 100 bp
 * 15816
 * 16551: contig of 736 bp in length
 * 16552
 * 16551: gap of 100 bp
 * 16552
 * 17384: contig of 733 bp in length
 * 17385
 * 17484: gap of 100 bp
 * 17485
 * 18195: contig of 711 bp in length
 * 18196
 * 18295: gap of 100 bp
 * 18296
 * 19018: contig of 723 bp in length
 * 19019
 * 19118: gap of 100 bp
 * 19119
 * 19544: contig of 736 bp in length
 * 19545
 * 19544: gap of 100 bp
 * 19555
 * 20702: contig of 748 bp in length
 * 20703
 * 20802: gap of 100 bp
 * 20803
 * 21524: contig of 722 bp in length
 * 21525
 * 21624: gap of 100 bp
 * 21625
 * 22364: contig of 740 bp in length
 * 22365
 * 22464: gap of 100 bp
 * 22465
 * 23180: contig of 716 bp in length
 * 23181
 * 23280: gap of 100 bp
 * 23281
 * 23991: contig of 711 bp in length
 * 23992
 * 24091: gap of 100 bp
 * 24092
 * 24817: contig of 726 bp in length
 * 24818
 * 24917: gap of 100 bp
 * 24918
 * 25648: contig of 731 bp in length
 * 25649
 * 25748: gap of 100 bp
 * 25749
 * 26489: contig of 741 bp in length
 * 26490
 * 26589: gap of 100 bp
 * 26590
 * 27323: contig of 734 bp in length
 * 27324
 * 27423: gap of 100 bp
 * 27424
 * 28169: contig of 746 bp in length
 * 28170
 * 28269: gap of 100 bp
 * 28270
 * 29006: contig of 737 bp in length
 * 29106
 * 29106: gap of 100 bp
 * 29107
 * 29850: contig of 744 bp in length
 * 29851
 * 29950: gap of 100 bp
 * 29951
 * 30687: contig of 737 bp in length
 * 30688
 * 30787: gap of 100 bp
 * 30788
 * 31505: contig of 718 bp in length
 * 31506
 * 31605: gap of 100 bp
 * 31606
 * 32300: contig of 695 bp in length
 * 32301
 * 32400: gap of 100 bp
 * 32401
 * 33111: contig of 711 bp in length
 * 33112
 * 33211: gap of 100 bp
 * 33212
 * 33931: contig of 720 bp in length
 * 33932
 * 34031: gap of 100 bp
 * 34032
 * 34769: contig of 738 bp in length
 * 34770
 * 34869: gap of 100 bp
 * 34870
 * 35596: contig of 727 bp in length
 * 35597
 * 35696: gap of 100 bp
 * 35697
 * 36433: contig of 737 bp in length
 * 36434
 * 36533: gap of 100 bp
 * 36534
 * 37249: contig of 716 bp in length
 * 37250
 * 37349: gap of 100 bp
 * 37350
 * 38078: contig of 729 bp in length
 * 38079
 * 38178: gap of 100 bp
 * 38179
 * 38930: contig of 752 bp in length
 * 38931
 * 39030: gap of 100 bp
 * 39031
 * 39780: contig of 750 bp in length
 * 39781
 * 39880: gap of 100 bp
 * 39881
 * 40614: contig of 734 bp in length
 * 40615
 * 40714: gap of 100 bp
 * 40715
 * 41446: contig of 732 bp in length
 * 41447
 * 41546: gap of 100 bp
 * 41547
 * 42250: contig of 704 bp in length
 * 42251
 * 42350: gap of 100 bp
 * 42351
 * 43096: contig of 746 bp in length
 * 43097
 * 43196: gap of 100 bp

* 43197	43925:	contig of 729 bp	in length
* 43926	44025:	gap of 100 bp	
* 44026	44755:	contig of 730 bp	in length
* 44756	44855:	gap of 100 bp	
* 44856	45882:	contig of 727 bp	in length
* 45883	45682:	gap of 100 bp	
* 45683	46425:	contig of 743 bp	in length
* 46426	46525:	gap of 100 bp	
* 46526	47272:	contig of 747 bp	in length
* 47273	47372:	gap of 100 bp	
* 47373	48117:	contig of 745 bp	in length
* 48118	48217:	gap of 100 bp	
* 48218	48962:	contig of 745 bp	in length
* 48963	49062:	gap of 100 bp	
* 49063	49768:	contig of 706 bp	in length
* 49769	49868:	gap of 100 bp	
* 49869	50606:	contig of 738 bp	in length
* 50607	50706:	gap of 100 bp	
* 50707	51435:	contig of 729 bp	in length
* 51436	51535:	gap of 100 bp	
* 51536	52281:	contig of 746 bp	in length
* 52282	52381:	gap of 100 bp	
* 53126	53225:	contig of 744 bp	in length
* 53226	53975:	contig of 750 bp	in length
* 53976	54075:	gap of 100 bp	
* 54076	54782:	contig of 707 bp	in length
* 54783	54882:	gap of 100 bp	
* 54883	55623:	contig of 741 bp	in length
* 55624	55723:	gap of 100 bp	
* 55724	56456:	contig of 733 bp	in length
* 56457	56556:	gap of 100 bp	
* 57284	57284:	contig of 728 bp	in length
* 57285	57384:	gap of 100 bp	
Alignment Scores:			
Pred. No.:	5.11e+03	Length:	72957
Score:	43.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	89.6%	Indels:	0
DB:	14	Gaps:	0
US-10-774-176-9 (1-9) x AC025304 (1-72957)			
Qy 1 GlyLeuProHisIleArgValPheLeu 9			
Db 48630 GGACTCCCTCACTTGCGTTTGTCCTG 48604			
RESULT 26			
AC003062			
LOCUS			
DEFINITION Mouse Chromosome 16 Region Syntenic to DSCR BAC Clone b264n1, linear ROD 18-JAN-2002			
ACCESSION AC003062			
VERSION AC003062.2 GI:4731672			
KEYWORDS HTG.			
SOURCE Mus musculus (house mouse)			
ORGANISM Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.			
REFERENCE 1 (bases 1 to 89743)			
AUTHORS Chen, F. and Roe, B.A.			
JOURNAL Unpublished			
REFERENCE 2 (bases 1 to 89743)			
AUTHORS Galili, N., Baldwin, S., Lund, J., Reeves, R., Gong, W., Chen, F., Roe, B.A., Emanuel, B.S., Nayak, S., Mickanin, C., Budarf, M.L. and Buck, C.A.			
TITLE A Region of Mouse Chromosome 16 is Syntenic to the DiGeorge, Velo-Cardio-Facial Syndrome Minimal Critical Region			
JOURNAL Unpublished			
REMARK The genes were identified by comparing with human genomic and cDNA			
sequences and RT-PCR of 12 day post conception mouse embryos total RNA			
3 (bases 1 to 89743)			
Roe, B.A. Dr.			
Direct Submission			
Submitted (06-NOV-1997) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval Room 208, Norman, OK 73019, USA			
4 (bases 1 to 89743)			
Roe, B.A. Dr.			
Direct Submission			
Submitted (16-SEP-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval Room 208, Norman, OK 73019, USA			
5 (bases 1 to 89743)			
Roe, B.A. Dr.			
Direct Submission			
Submitted (23-SEP-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval Room 208, Norman, OK 73019, USA			
6 (bases 1 to 89743)			
Roe, B.A. Dr.			
Direct Submission			
Submitted (05-NOV-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval Room 208, Norman, OK 73019, USA			
7 (bases 1 to 89743)			
Roe, B.A. Dr.			
Direct Submission			
Submitted (03-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval Room 208, Norman, OK 73019, USA			
8 (bases 1 to 89743)			
Roe, B.A. Dr.			
Direct Submission			
Submitted (15-FEB-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval Room 208, Norman, OK 73019, USA			
9 (bases 1 to 89743)			
Roe, B.A. Dr.			
Direct Submission			
Submitted (18-JAN-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval Room 208, Norman, OK 73019, USA			
On May 3, 1999 this sequence version replaced gi:3845373.			
Center: Department Of Chemistry And Biochemistry			
The University Of Oklahoma			
Center code: UOKNOR			
FEATURES			
Location/Qualifiers			
1..89743			
/organism="Mus musculus"			
/mol_type="genomic DNA"			
/strain="129/Sv EV cell line CJ7"			
/db_xref="taxon:10090"			
/chromosome="16"			
/clone="ct7-264n1"			
/clone_lib="CitbCJ7 mouse BAC library"			
ORIGIN			
Alignment Scores:			
Pred. No.:	6.34e+03	Length:	89743
Score:	43.00	Matches:	8
Percent Similarity:	88.9%	Conservative:	0
Best Local Similarity:	88.9%	Mismatches:	1
Query Match:	89.6%	Indels:	0
DB:	9	Gaps:	0
US-10-774-176-9 (1-9) x AC003062 (1-89743)			
Qy 1 GlyLeuProHisIleArgValPheLeu 9			

Db	56692	GGGTTACTCATATTAGATATTCTTA	56718	
RESULT 27	AE013598_20/c			
WFCOMMENT	Sequence split into 50 fragments	LOCUS AE013598	Accession AE013598	
	Fragment Name	Begin	End	
	AE013598_00	1	110000	
	AE013598_01	100001	210000	
	AE013598_02	200001	310000	
	AE013598_03	300001	410000	
	AE013598_04	400001	510000	
	AE013598_05	500001	610000	
	AE013598_06	600001	710000	
	AE013598_07	700001	810000	
	AE013598_08	800001	910000	
	AE013598_09	900001	1010000	
	AE013598_10	1000001	1110000	
	AE013598_11	1100001	1210000	
	AE013598_12	1200001	1310000	
	AE013598_13	1300001	1410000	
	AE013598_14	1400001	1510000	
	AE013598_15	1500001	1610000	
	AE013598_16	1600001	1710000	
	AE013598_17	1700001	1810000	
	AE013598_18	1800001	1910000	
	AE013598_19	1900001	2010000	
	AE013598_20	2000001	2110000	
	AE013598_21	2100001	2210000	
	AE013598_22	2200001	2310000	
	AE013598_23	2300001	2410000	
	AE013598_24	2400001	2510000	
	AE013598_25	2500001	2610000	
	AE013598_26	2600001	2710000	
	AE013598_27	2700001	2810000	
	AE013598_28	2800001	2910000	
	AE013598_29	2900001	3010000	
	AE013598_30	3000001	3110000	
	AE013598_31	3100001	3210000	
	AE013598_32	3200001	3310000	
	AE013598_33	3300001	3410000	
	AE013598_34	3400001	3510000	
	AE013598_35	3500001	3610000	
	AE013598_36	3600001	3710000	
	AE013598_37	3700001	3810000	
	AE013598_38	3800001	3910000	
	AE013598_39	3900001	4010000	
	AE013598_40	4000001	4110000	
	AE013598_41	4100001	4210000	
	AE013598_42	4200001	4310000	
	AE013598_43	4300001	4410000	
	AE013598_44	4400001	4510000	
	AE013598_45	4500001	4610000	
	AE013598_46	4600001	4710000	
	AE013598_47	4700001	4810000	
	AE013598_48	4800001	4910000	
	AE013598_49	4900001	4941439	
	Continuation (21 of 50) of AE013598 from base 2000001 (AE013598 oryzae pv. c			
Alignment Scores:				
Pred. No.:	7.84e+03	Length:	110000	
Score:	43.00	Matches:	7	
Percent Similarity:	88.9%	Conservative:	1	
Best Local Similarity:	77.8%	Mismatches:	1	
Query Match:	89.6%	Indels:	0	
DB:	1	Gaps:	0	
US-10-774-176-9 (1-9) x AE013598_20 (1-110000)				
Qy	1	GlyLeuProHisIleArgValPheLeu	9	
Db	76510	GGCCCTCCCGCCAGTCGGAACCTTCTTA	76484	
RESULT 28	AC005816			
LOCUS	AC005816			
DEFINITION	Mus musculus strain 129/Sv clone ct7-326b16 map 16, complete sequence.			
ACCESSION	AC005816			
VERSION	AC005816.6			
KEYWORDS	HTG.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Mus musculus Chromosome 16 BAC Clone ct7-326b16 In DGC Region Unpublished			
JOURNAL	2 (bases 1 to 136687)			
REFERENCE	2 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-OCT-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	3 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (10-DEC-1998) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	4 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (03-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	5 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (04-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	6 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (06-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	7 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-FEB-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	8 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-FEB-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	9 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-JAN-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	10 (bases 1 to 136687)			
AUTHORS	Chen, F., Do, T., Reeves, R.H. and Roe, B.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-SEP-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
COMMENT	On Sep 17, 2003 this sequence version replaced gi:7019299.			
	----- Genome Center			

Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES

source
Location/Qualifiers
1..136687
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/map="16"
/clone="ct7-326b16"
/clone_lib="Citbcu7 mouse BAC library"

ORIGIN

Alignment Scores:
Pred. No.: 9.83e+03 Length: 136687
Score: 43.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.6% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-9 (1-9) x AC005816 (1-136687)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 135667 GGGTTACCTCATATTAGATATTCTTA 135693

RESULT 29
AC079831/c 147412 bp DNA linear ROD 13-FEB-2003
LOCUS Mus musculus strain C57BL/6J chromosome 16 clone rp23-374p12,
DEFINITION complete sequence.

ACCESSION AC079831
VERSION AC079831.24 GI:24961478
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Unpublished
TITL Mus musculus BAC Clone rp23-374p12

REFERENCE 2 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Direct Submission
TITL Direct Submission

REFERENCE 3 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Submitted (12-SEP-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 4 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Direct Submission
TITL Direct Submission

REFERENCE 5 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Submitted (01-AUG-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 6 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Direct Submission
TITL Direct Submission

REFERENCE 7 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Submitted (14-NOV-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 8 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Direct Submission
TITL Direct Submission

REFERENCE 9 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 10 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
JOURNAL Direct Submission
TITL Direct Submission

JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

Submitted (08-JAN-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

7 (bases 1 to 147412)
Song, L., Jiang, X. and Roe, B.A.
Direct Submission

Submitted (13-FEB-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

On Nov 14, 2002 this sequence version replaced gi:22038582.

Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

FEATURES

source
Location/Qualifiers
1..147412
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="16"
/clone="rp23-374p12"
/clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"

ORIGIN

Alignment Scores:
Pred. No.: 1.06e+04 Length: 147412
Score: 43.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.6% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-9 (1-9) x AC079831 (1-147412)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 116229 GGGTTACCTCATATTAGATATTCTTA 116203

RESULT 30
AC135216/c

LOCUS Bos taurus clone RP42-427020, WORKING DRAFT SEQUENCE, 3 ordered
DEFINITION pieces.
ACCESSION AC135216
VERSION AC135216.2 GI:25699761
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 153264)
Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.O., Brinkley, C., Brooks, S.,
Carla, C., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Maguiness, E.H., Masiello, C., Maskeri, B., McDowell, J.,
Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
Reddix-Dugue, N., Schandler, K., Schuster, M.G., Sison, C.,
Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

REFERENCE 2 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 3 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 4 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 5 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 6 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 7 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 8 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 9 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

REFERENCE 10 (bases 1 to 153264)
Green, E.D.
Direct Submission
Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovesmont Circle, Gaithersburg, MD 20877, USA

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Nov 27, 2002 this sequence version replaced gi:23622199.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: djz
Center clone name: 427020

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 152954 bases at least Q40
Consensus quality: 153013 bases at least Q30
Consensus quality: 153042 bases at least Q20
Insert size: 140000; agarose-fp
Insert size: 153064; sum-of-contigs
Quality coverage: 11.86x in Q20 bases; agarose-fp
Quality coverage: 10.84x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 69854: contig of 69854 bp in length
* 69855 108944: contig of 38990 bp in length
* 108945 109044: gap of unknown length
* 109045 153264: contig of 44220 bp in length.

FEATURES

Location/Qualifiers
1..153264
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-427020"
/clone_lib="RP42"

misc_feature
1..69854
/note="assembly_fragment
clone end:77
vector_side:left"
69855..69954
/estimated_length=unknown
69955..108944
/note="assembly_fragment"
108945..109044
/estimated_length=unknown
109045..153264
/note="assembly_fragment
clone end:SP6
vector_side:right"

ORIGIN

Alignment Scores:
Pred. No.: 1.11e+04 Length: 153264
Score: 43.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.6% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-9 (1-9) x AC135216 (1-153264)
QY 1 GlyLeuProHisIleArgValPheLeu 9
|||||
Db 140629 GGGCTACCATACATACAGATTTCTG 140603
|||||

AC022876 158546 bp DNA linear HTG 01-MAR-2000
Homo sapiens chromosome 11 clone RP11-398P9 map 11, WORKING DRAFT
SEQUENCE, 44 unordered pieces.
AC022876
AC022876.2 GI:7139773
HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 (bases 1 to 158546)
AUTHORS
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE
Homo sapiens chromosome 11, clone RP11-398P9
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 158546)
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Buckett,G., Castle,A., Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArlelano,K., Dewar,K., Domino,M., Doyle,M., Fenesstor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McKwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
TITLE
JOURNAL
COMMENT
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 1, 2000 this sequence version replaced gi:6922086.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6209
Center clone name: 398_P_9
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 136264 bases at least Q40
Consensus quality: 146380 bases at least Q30
Consensus quality: 150509 bases at least Q20
Insert size: 170000; agarose-fp


```

Pred. No.: 1.15e+04 Length: 158546
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.6% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-9 (1-9) x AC022876 (1-158546)

OY 1 GlyLeuProHisIleArgValPheLeu 9
Db 94101 GGGCTGCCTCATATACGTGTATATA 94127

RESULT 32
AC009949 AC009949 174699 bp DNA linear PRI 15-APR-2005
LOCUS Homo sapiens BAC clone RP11-69J7 from 2, complete sequence.
DEFINITION AC009949
ACCESSION AC009949
VERSION AC009949.9 GI:11386326
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo
REFERENCE
  1 (bases 1 to 174699)
  Abbott.A., McLeellan.M. and Haub,K.
  The sequence of Homo sapiens BAC clone RP11-69J7
  Unpublished (2001)
REFERENCE
  2 (bases 1 to 174699)
  Waterston,R.H.
  Direct Submission
  Submitted (08-SEP-1999) Genome Sequencing Center, Washington
  University School of Medicine, 4444 Forest Park Parkway, St. Louis,
  MO 63108, USA
REFERENCE
  3 (bases 1 to 174699)
  Waterston,R.H.
  Direct Submission
  Submitted (28-NOV-2000) Genome Sequencing Center, Washington
  University School of Medicine, 4444 Forest Park Parkway, St. Louis,
  MO 63108, USA
REFERENCE
  4 (bases 1 to 174699)
  Waterston,R.H.
  Direct Submission
  Submitted (09-MAY-2001) Department of Genetics, Washington
  University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
  5 (bases 1 to 174699)
  Waterston,R.H.
  Direct Submission
  Submitted (03-JAN-2002) Genome Sequencing Center, Washington
  University School of Medicine, 4444 Forest Park Parkway, St. Louis,
  MO 63108, USA
REFERENCE
  6 (bases 1 to 174699)
  Waterston,R.H.
  Direct Submission
  Submitted (29-OCT-2002) Department of Genetics, Washington
  University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
  7 (bases 1 to 174699)
  Wilson,R.K.
  Direct Submission
  Submitted (15-APR-2005) Genome Sequencing Center, Washington
  University School of Medicine, 4444 Forest Park Parkway, St. Louis,
  MO 63108, USA
COMMENT
  On Nov 28, 2000 this sequence version replaced gi:10337664.
  -----
  Center: Genome Center
  Center: Washington University Genome Sequencing Center
  Center code: WUGSC
  Web site: http://genome.wustl.edu
  Contact: submissions@watson.wustl.edu
  -----
  Summary Statistics
  Center project name: H_NH0069J07
  -----

```

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phased quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanesi,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pRACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-98F19; the clone sequenced to the right is RP11-395A23. Actual start of this clone is at base position 1 of RP11-69J7; actual end is at base position 174699 of RP11-69J7.

```

FEATURES
    source
        1. .174699
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="2"
            /clone="RP11-69J7"
            /clone_lib="RP11-11"
            4511. .48194
            /gene="SP140"
            join(4511. .4597,4816. .4932,5565. .5618,19325. .19390,
20731. .20811,22871. .22945,25439. .25543,27625. .27667,
29250. .29297,32400. .32441,44903. .45018,45722. .45830,
46133. .46210,46431. .46574,47565. .48194)
            /gene="SP140"
            join(<4511. .4597,4816. .4932,5565. .5618,19325. .19390,
20731. .20811,22871. .22945,25439. .25543,27625. .27667,
29250. .29297,32400. .32441,44903. .45018,45722. .45830,
46133. .46210,46431. .46574,47565. .47663)
            /gene="SP140"
            /note="Homo sapiens SP140 nuclear body protein (SP140),
mRNA.; H_NH0069J07.1
This gene was based on gi(21361373)
Continued from H_NH0098F19.2"
            /codon_start=3
            /product="unknown"
            /protein_id="AA88868.1"
            /db_xref="GI:62630122"
            /translation="SSELENHPNNEGESEELASSLLYDNPQAEQAYENKCSVM
CFSEEVPGSPFEARTSDQACGTVDIANNSTLGPKRKRKRKGHSGMRMRQRE
NQNDNSKADQGVSSKKNVNLKDLIRGRKRGKPGTRFTQSDAAQKRVRSRA
SKHKDETVDKAPLLPVTCGGVKGILHKKLQGLVKICIQTEDGKFWPTPEIKG
GHARKNWLISVRCGFWPLWLMENGFLDPDPPIRYRKKRILKSONNSVDPQMRNL
DCEVCRDGEGLFCDDTSKRVFHEDCHIPPVEARIPWNCIFMKKSSPSQOCCQES
EVLREQMCPPEQLKCEFLLLKLVYCCSESSFAPKIPYYIIRACQGLKBPMLDKIKK
RLNEHGYPQVEGFQDMRLIFQNHRSAYKYKDFQGMGRLEAFKFNKFEVPAIQETN
GNN"
            77005. .77343
            /note="CpG island (AGC=59.6, o/e=0.71, #CpGs=25)"
            105885. .138709
            /gene="LOC93349"
            join(105885. .105946,106577. .106630,118489. .118554,

```

```

misc_feature
gene
mRNA

```


Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GTDP
 Center clone name: CH230-391M15
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 157421 bases at least Q40
 Consensus quality: 161044 bases at least Q30
 Consensus quality: 163069 bases at least Q20
 Estimated insert size: 157850; sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 181686: contig of 181686 bp in length
 * 181687 181786: gap of unknown length
 * 181787 183066: contig of 1280 bp in length.

FEATURES

source
 Location/Qualifiers
 1..183066
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-391M15"
 misc_feature
 1..800
 /note="clone_boundary
 clone_end:Sp6
 site:
 end_sequence:BZ209492"
 118887..120169
 /note="wgs_contig"
 misc_feature
 120220..121383
 /note="wgs_contig"
 misc_feature
 151817..153060
 /note="wgs_contig"
 misc_feature
 167845..170206
 /note="wgs_contig"
 misc_feature
 178674..179842
 /note="wgs_contig"
 misc_feature
 complement[180647..181536]
 /note="clone_boundary
 clone_end:17
 site:
 end_sequence:BZ209490"
 181687..181786
 /estimated_length-unknown

gap

ORIGIN

Alignment Scores:
 Pred. No.: 1.33e+04 Length: 183066
 Score: 43.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 89.6% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-9 (1-9) x AC123238 (1-183066)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
 |||:::|||||:::|||||
 Db 3635 GGTGTACCCACCTAAGGCTCTCTTA 3609

RESULT 34

AC008020 AC008020 184044 bp DNA linear ROD 19-OCT-2002
 LOCUS
 DEFINITION Mus musculus strain 129S6/SvEvTac chromosome 16 clone rp21-598k13,

complete sequence.

AC008020
 VERSION
 AC008020.37 GI:24137494
 KEYWORDS
 HTG.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Mus musculus Chromosome 16 PAC Clone rp21-598k13

JOURNAL Unpublished

REFERENCE
 AUTHORS Deschamps, S., Oomen, S. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-1999)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2000)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (03-MAY-2000)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (05-MAY-2000)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (11-MAY-2000)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (18-MAY-2000)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (11-JUL-2000)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (15-SEP-2000)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (29-JAN-2002)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE
 AUTHORS Deschamps, S., Oomen, S., Draber, R., Becat, C. and Roe, B.A.

TITLE Direct Submission

JOURNAL Submitted (09-JUL-2002)

Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 12 (bases 1 to 184044)
 AUTHORS Deschamps,S., Omen,S., Draber,R., Becat,C. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 13 (bases 1 to 184044)
 AUTHORS Deschamps,S., Omen,S., Draber,R., Becat,C. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (25-AUG-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE 14 (bases 1 to 184044)
 AUTHORS Deschamps,S., Omen,S., Draber,R., Becat,C. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (19-OCT-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

COMMENT
 OK 73019, USA
 On Oct 19, 2002 this sequence version replaced gi:22475337.
 Center: Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKNOR

FEATURES
 source
 Location/Qualifiers
 1..184044
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129S6/SvEvTac"
 /db_xref="taxon:10090"
 /chromosome="16"
 /clone_lib="RPC1 - 21 Female (129S6/SvEvTac) Mouse PAC
 Library"

ORIGIN
 Alignment Scores: Length: 184044
 Pred. No.: 1.34e+04 Matches: 8
 Score: 43.00 Conservative: 0
 Percent Similarity: 88.9% Mismatches: 1
 Best Local Similarity: 88.9% Indels: 0
 Query Match: 89.6% Gaps: 0
 DB: 9

US-10-774-176-9 (1-9) x AC008020 (1-184044)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
 Db 31951 GGGTTACCTCATATTAGATATTTCTTA 31977

RESULT 35
 AP002353/c
 LOCUS Homo sapiens genomic DNA, chromosome 11 clone:CTD-2651C21, complete
 DEFINITION sequence.
 ACCESSION AP002353
 VERSION AP002353.4 GI:31790676
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE 1
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA
 JOURNAL Published Only in Database (2000)
 REFERENCE 2 (bases 1 to 202607)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (25-MAY-2000) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 On Jun 16, 2003 this sequence version replaced gi:19263027.

COMMENT
 On Jun 16, 2003 this sequence version replaced gi:19263027.

FEATURES
 source
 Location/Qualifiers
 1..202607
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="CTD-2651C21"

ORIGIN
 Alignment Scores: Length: 202607
 Pred. No.: 1.48e+04 Matches: 7
 Score: 43.00 Conservative: 2
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 77.8% Indels: 0
 Query Match: 89.6% Gaps: 0
 DB: 8

US-10-774-176-9 (1-9) x AP002353 (1-202607)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
 Db 15010 GGGCTGCTCATATACGGTGTATATA 14984

RESULT 36
 AC102169/c
 LOCUS Mus musculus chromosome 10 clone RP23-455E13 map 10, *** SEQUENCING
 DEFINITION IN PROGRESS ***, 12 unordered pieces.
 ACCESSION AC102169
 VERSION AC102169.3 GI:42628089
 KEYWORDS HTG; HTGS PHASE1; HTGS FULLTOP; HTGS ACTIVEPIN.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 204343)
 AUTHORS Birren,B., Nusbaum,C. and Lander,E.
 TITLE Mus musculus chromosome 10, clone RP23-455E13
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 204343)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., DeLorellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyne,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Katat,A., Karatas,A., Kells,C., LaRoque,K.,
 Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Riese,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted Submission
 Direct Submission
 Title
 JOURNAL
 REFERENCE 3 (bases 1 to 204343)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
 Boguslavsky,L., Boukhalter,B., Canarata,J., Chang,J.-J., Choepel,Y.,
 Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
 Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrum,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

Submitted (19-FEB-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Feb 19, 2004 this sequence version replaced gi:22004590.
 All repeats were identified using RepeatMasker:

Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L18160

Center clone name: 455_E_13

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available with the accession number will
 * be preserved.

* 1 34390: contig of 34390 bp in length
 * 34391 34490: gap of 100 bp
 * 34491 43838: contig of 9348 bp in length
 * 43839 43938: gap of 100 bp
 * 43939 57302: contig of 13364 bp in length
 * 57303 57402: gap of 100 bp
 * 57403 61947: contig of 4545 bp in length
 * 61948 62047: gap of 100 bp
 * 62048 86421: contig of 24374 bp in length
 * 86422 86521: gap of 100 bp
 * 86522 95722: contig of 9201 bp in length
 * 95723 106208: contig of 10386 bp in length
 * 106209 106308: gap of 100 bp
 * 106309 138441: contig of 32133 bp in length
 * 138442 138541: gap of 100 bp
 * 138542 143561: contig of 5020 bp in length
 * 143562 143661: gap of 100 bp
 * 143662 153566: contig of 9905 bp in length
 * 153567 153666: gap of 100 bp
 * 153667 183508: contig of 29842 bp in length
 * 183509 183608: gap of 100 bp
 * 183609 204343: contig of 20735 bp in length.

FEATURES

source

1..204343
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="10"

/map="10"
 /clone="RP23-455E13"
 /clone_lib="RP23-23 Female Mouse BAC"

gap 34391..34490
 /estimated_length=100
 gap 43839..43938
 /estimated_length=100
 gap 57303..57402
 /estimated_length=100
 gap 61948..62047
 /estimated_length=100
 gap 86422..86521
 /estimated_length=100
 gap 95723..95822
 /estimated_length=100
 gap 106209..106308
 /estimated_length=100
 gap 138442..138541
 /estimated_length=100
 gap 143562..143661
 /estimated_length=100
 gap 153567..153666
 /estimated_length=100
 gap 183509..183608
 /estimated_length=100

ORIGIN

Alignment Scores:
 Pred. No.: 1..49e+04 Length: 204343
 Score: 43.00 Matches: 7
 Percent Similarity: 100.0% Conservatives: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 89.6% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-9 (1-9) x AC102169 (1-204343)

QY 1 GlyLeuProHisIleArgValPheLeu 9

Db 163476 GGGGTCCCACTGTGAGAGTATTTTAA 163450

RESULT 37

AC166253

LOCUS AC166253 212065 bp DNA linear HTG 31-JUL-2005
 DEFINITION Mus musculus clone RP23-231M6, WORKING DRAFT SEQUENCE, 12 unordered
 pieces.

AC166253

AC166253.3 GI:71534302

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.

Muzny,D., Adams,C., Agbai II,O., Allen,C., Alsbrooks,S., Archer,P.,

Arredondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R.,

Berarducci,A., Biswalo,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,

Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,

Chacko,J., Chahrour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,

Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,

Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De

Anda,C., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H.,

Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,

Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,

Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,

Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,

Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,

Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haebleren,K.,

Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,

Hawes,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,

Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,B.,

Howell,L.T., Hulyk,S., Hume,J., Imo,K., Jackson,A., Jackson,L.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaika, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, C., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Poster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorogis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, R., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louleeged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawney, S., McLeod, M. P., McNeill, T. Z., Meenen, S., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfanckoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Fu, L. L., Puzo, M., Racho, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, S., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (01-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jun 29, 2005 this sequence version replaced gi:66792995.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FH2

Center clone name: CH240-121K5

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 170996 bases at least Q40

Consensus quality: 192270 bases at least Q30

Consensus quality: 201398 bases at least Q20

Estimated insert size: 240476; sum-of-contigs estimation

Estimated insert size: 103626; agarose-fp estimation

Quality coverage: 7x in Q20 bases; agarose-fp estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

consists of 55 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

* 1 4570: contig of 4570 bp in length

* 4571 4620: gap of 50 bp

* 4621 9936: contig of 5216 bp in length

* 9937 15829: contig of 5893 bp in length

* 15830 18073: contig of 2194 bp in length

* 18074 18173: gap of unknown length

* 18174 22230: gap of 50 bp

* 22231 28801: contig of 6571 bp in length

* 28802 30422: contig of 1521 bp in length

* 30423 31731: gap of 1309 bp

* 31732 33350: contig of 1619 bp in length

* 33351 46349: contig of 12899 bp in length

* 46350 53109: contig of 6710 bp in length

* 53110 56978: contig of 3769 bp in length

* 56979 62347: contig of 5318 bp in length

* 62348 63449: contig of 1003 bp in length

* 63450 65048: contig of 1499 bp in length

* 65049 66550: contig of 1502 bp in length

* 66551 68385: contig of 1635 bp in length

* 68386 69554: contig of 1069 bp in length

* 69555 71570: contig of 1916 bp in length

* 71571 72928: contig of 1158 bp in length

* 72929 74494: contig of 1566 bp in length

* 74495 76037: contig of 1442 bp in length

* 76038 77443: contig of 1307 bp in length

* 77444 77443: contig of 1307 bp in length

```

* 77444 77543: gap of unknown length
* 77544 78654: contig of 1111 bp in length
* 78655 78754: gap of unknown length
* 78755 80745: contig of 1991 bp in length
* 80746 82293: contig of 1448 bp in length
* 82294 82393: gap of unknown length
* 82394 84218: contig of 1725 bp in length
* 84219 87067: contig of 2849 bp in length
* 87068 90298: contig of 3131 bp in length
* 90299 93065: contig of 2667 bp in length
* 93066 93165: gap of unknown length
* 93166 94905: contig of 1740 bp in length
* 94906 96679: contig of 1674 bp in length
* 96680 97799: gap of unknown length
* 97799 99099: contig of 2320 bp in length
* 99100 101283: contig of 2084 bp in length
* 101284 104893: contig of 3510 bp in length
* 104894 107673: contig of 2680 bp in length
* 107674 110773: gap of unknown length
* 110774 110782: contig of 3009 bp in length
* 110783 113848: contig of 2966 bp in length
* 113849 117922: contig of 3974 bp in length
* 117923 122501: contig of 4479 bp in length
* 122502 126946: contig of 4345 bp in length
* 126947 130053: contig of 3007 bp in length
* 130054 132083: contig of 1930 bp in length
* 132084 132183: gap of unknown length

```

Alignment Scores:

Pred. No.:	1.58e+04	Length:	215042
Score:	43.00	Matches:	8
Percent Similarity:	88.9%	Conservative:	0
Best Local Similarity:	88.9%	Mismatches:	1
Query Match:	89.6%	Indels:	0
DB:	14	Gaps:	0

US-10-774-176-9 (1-9) x AC162395 (1-215042)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

Db 71402 GGACTGCCACATGCCAGGTTTCTG 71428

RESULT 39
AC133670/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-230811, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
AC133670 221701 bp DNA linear HTG 20-NOV-2002

AC133670
VERSION AC133670.2 GI:25139327
KEYWORDS HTG, HTGS PHASE1, HTGS DRAFT, HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Rattus
1 (bases 1 to 221701)
Muzny, D. Marie., Metker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Derman, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garza, M., Garza, M.,
Georgiadis, E., Geer, K., Gill, R., Grady, M., Guerra, T., Guera, W.,
Gunatne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. I., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, J., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, H., L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwackemeh, O., Okwono, G., Olarnpunaagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poinexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, E.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Soza, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 221701)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (17-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 221701)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KBFW
Center clone name: CH230-230B11
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 213982 bases at least Q40
Consensus quality: 215530 bases at least Q30
Consensus quality: 216707 bases at least Q20
Estimated insert size: 220267; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 2 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
* 1 220431: contig of 220431 bp in length
* 220432 220531: gap of unknown length
* 220532 221701: contig of 1170 bp in length.
-----
FEATURES             Location/Qualifiers
     source            1..221701
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-230B11"
     misc_feature      1..1905
                        /note="wgs contig"
     misc_feature      218601..220431
                        /note="wgs contig"
     gap               220432..220531
                        /estimated_length=unknown
ORIGIN
Alignment Scores:
Pred. No.: 1.63e+04 Length: 221701
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.6% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-9 (1-9) x AC133670 (1-221701)

QY 1 GlyLeuProHisIleArgValPheLeu 9
|||||:|||||:|||||:|||||:|||||
DB 87200 GGGCTTCCTCACTGAGACTTTTCCTT 87174

RESULT 40
AC113264
LOCUS AC113264 222589 bp DNA linear ROD 16-NOV-2002
DEFINITION Mus musculus strain C57BL/6J chromosome 16 clone rp23-357j7,
complete sequence.
ACCESSION AC113264
VERSION AC113264.21 GI:25046401
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 222589)
Deschamps, S., Li, Y., Hu, X. and Roe, B.A.
AUTHORS
TITLE
JOURNAL
REFERENCE
Mus musculus rp23-357j7

```

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 222589)
AUTHORS Li, Y., Hu, X. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 222589)
AUTHORS Li, Y., Hu, X. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (10-OCT-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 4 (bases 1 to 222589)
AUTHORS Li, Y., Hu, X. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 5 (bases 1 to 222589)
AUTHORS Deschamps, S., Li, Y., Hu, X. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2002) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Nov 16, 2002 this sequence version replaced gi:23957635.
-----
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR
-----
FEATURES             Location/Qualifiers
     source            1..222589
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /chromosome="16"
                        /clone="rp23-357j7"
                        /clone_lib="RPCL - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN
Alignment Scores:
Pred. No.: 1.63e+04 Length: 222589
Score: 43.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 89.6% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-9 (1-9) x AC113264 (1-222589)

QY 1 GlyLeuProHisIleArgValPheLeu 9
|||||:|||||:|||||:|||||:|||||
DB 150725 GGGTTACCTCATATTAGATATTTCTTA 150751

RESULT 41
AC133573/c
LOCUS AC133573 224138 bp DNA linear ROD 20-NOV-2002
DEFINITION Mus musculus strain C57BL/6J clone rp23-139f17, complete sequence.
ACCESSION AC133573
VERSION AC133573.11 GI:25139951
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 224138)
Deschamps, S. and Roe, B.A.
AUTHORS
TITLE
JOURNAL
REFERENCE
Mus musculus BAC clone rp23-139f17
Unpublished
2 (bases 1 to 224138)

```

AUTHORS
TITLE
JOURNAL
 Deschamps, S. and Roe, B.A.
 Direct Submission
 Submitted (14-SEP-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
 OK 73019, USA
REFERENCE
 3 (bases 1 to 224138)
AUTHORS
 Deschamps, S. and Roe, B.A.
TITLE
JOURNAL
 Submitted (16-NOV-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
 OK 73019, USA
REFERENCE
 4 (bases 1 to 224138)
AUTHORS
 Deschamps, S. and Roe, B.A.
TITLE
JOURNAL
 Submitted (20-NOV-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
 OK 73019, USA
COMMENT
 On Nov 20, 2002 this sequence version replaced gi:25046357.
 ----- Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKNOR

FEATURES
 source
 Location/Qualifiers
 1..224138
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="ip23-139f17"
 /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN
 Alignment Scores:
 Pred. No.: 1.65e+04 Length: 224138
 Score: 43.00 Matches: 8
 Percent Similarity: 88.9% Conservative: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 89.6% Indels: 0
 DB: 9 Gaps: 0
 US-10-774-176-9 (1-9) x AC133573 (1-224138)
 Qy 1 GlyLeuProHisIleArgValPheIeu 9
 |||||
 Db 77101 GGGTACCTCATATTAGATATTCTTA 77075
 RESULT 42
 AC158595/c
 LOCUS
 DEFINITION Mus musculus clone RP24-216B14, WORKING DRAFT SEQUENCE, 5 unordered
 pieces.
 AC158595
 AC158595.10 GI:171361715
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 231951)
 Musny,D., Adams,C., Agbai II,O., Allen,C., Alsbrooks,S., Archer,P.,
 Arradondo,H., Bandaranaike,D., Bangura,L., Beltran,B., Beltran,R.,
 Beraducci,A., Biewalo,K., Blyth,P., Bonham,H., Buhay,C., Burch,P.,
 Cadoree,I., Canada,A., Cardenas,V., Carter,K., Cavazos,I.,
 Chacko,J., Chahour,M., Chavez,D., Chen,A., Chen,G., Chen,R.,
 Cheng,M.-T., Chu,J., Clerc,K., Cockrell,R., Coyle,M., Cree,A.,
 Curry,S., Dai,W., Davila,M.L., Davis,C., Davy-Carroll,L., De
 Anda,C., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H.,
 Donlin,J., McCauley,S., Dugan-Rocha,S., Dunn,A., Durbin,K.,
 Dziuda,D., Egan,A., Escotto,M., Espinosa,V., Eugene,C., Fa,M.,
 Fernandez,S., Fernando,P., Flagg,N., Forbes,L., Foster,P.,
 Fowler,G., Fu,Q., Fuh,E., Garcia,A., Garcia,R., Garner,T.,

Gaskin,C., Gench,S., Ghose,S., Gill,R., Gonzalez,D.,
 Gonzalez-Garay,M., Guevara,W., Holder,M., Haaland,W., Haerli,K.,
 Hall,B., Hamid,H., Hamilton,K., Harbes,B., Harris,R., Havlak,P.,
 Haves,A., Hawkins,E., Hayes,S., Hemphill,L., Hernandez,J.,
 Hines,S., Hitchens,M., Hodgson,A., Hogues,M., Hollins,J.,
 Howell,L.T., Hulyk,S., Hume,J., Ito,K., Jackson,A., Jackson,L.,
 Jacob,L., Jiang,H., Johnson,B., Johnson,R., Kalafus,K., Kelly,S.,
 Keys,T., Khan,Z., King,L., Kovar,C., Kowis,A., Kowis,C., Lara,F.,
 Leal,S., Lee,K., Lee,S., LeGall,F.I., Lemon,S., Lewis,L., Li,B.,
 Li,Y., Li,Z., Linnell,M., Liu,W., Liu,Y.-S., Liu,Y., Liyanage,D.,
 London,P., Lopez,J., Lorensunewa,L., Lozano,R., Luk,T., Madu,R.,
 Maheshwari,M., Mahoney,C., Malloy,K., Mansouri,D., Martinez,E.,
 McCallum,H., McPherson,J., Mercadao,C., Metzger,M.,
 Milosavljevic,A., Minja,E., Morgan,M., Morris,S., Munitadga,M.,
 Murray,D., Nazareth,L., Ngo,D., Nguyen,N., Norwig-Bastaugh,E.,
 Nott,A., Nwankwelu,O., Obregon,M., Ochi-Okorie,C., Odeh,E.,
 Okwuonu,G., Okwuonu,K., Parker,D., Pasternak,S., Patel,B.,
 Patel,Y., Paul,H., Perez,A., Perez,L., Petrosino,J., Pham,T.,
 Primus,E., Pu,L.-L., Puazo,M., Qin,X., Quinn,A., Quiroz,J.,
 Rabata,D., Rachlin,E., Reigh,R., Ren,Y., Reuter,M., Richards,S.,
 Rives,C., Rodriguez,P., Rojas,A., Ruiz,S.J., Sana,M., Sanders,W.,
 Santibanez,J., Santos,R., Savary,G., Scherer,S., Shen,H., Shen,Y.,
 Sison,I., Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R.,
 Svatek,A., Taylor,E., Taylor,T., Thomas,N., Thorn,R., Thornton,R.,
 Trejos,Z., Umani,K., Vargo,C., Verduzco,D., Villaseana,D., Virk,D.,
 Volkov,A., Waldron,L., Walker,B., Wang,O., Wang,S., Warren,J.,
 Wei,X., Wheeler,D., Williams,G., Williams,R., Worley,K., Wright,R.,
 Wu,J., Yakub,S., Yan,K., Yuan,Y., Yu,F., Zhang,J., Zhang,L.,
 Zhang,Z., Zhou,J., Weinstein,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 231951)
 Worley,K.C.
 Direct Submission
 Submitted (23-MAR-2005) Human Genome Sequencing Center, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 231951)
 Worley,K.C.
 Direct Submission
 Submitted (29-JUL-2005) Human Genome Sequencing Center, Baylor
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jul 28, 2005 this sequence version replaced gi:71274218.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help.tmc.edu
 ----- Project Information
 Center project name: MCGJR
 Center clone name: RP24-216B14
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Sequencing vector: M13;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 23253 bases at least Q40
 Consensus quality: 233090 bases at least Q30
 Consensus quality: 233474 bases at least Q20
 Estimated insert size: 248318; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 104177: contig of 104177 bp in length
 * 104178 104277: gap of unknown length

* as soon as it is available and the accession number will
* be preserved.

* 1 230429: contig of 230429 bp in length
* 230430 230529: gap of unknown length
* 230530 232670: contig of 2141 bp in length.

FEATURES

source Location/Qualifiers

1. 232670
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-102H13"
1. 11154
/note="wgs end extension
clone end:Sp6"
4426..5307
/note="clone boundary
clone end:Sp6"

misc_feature

4426..5307
/note="clone boundary
clone end:Sp6"

misc_feature

4426..5307
/note="clone boundary
clone end:Sp6"

misc_feature

end sequence:BH289845"
complement(225585..226409)
/note="clone boundary
clone end:T7"

misc_feature

site:EcORI
end sequence:BH289843"
228591..230429

misc_feature

/note="wgs end extension
clone end:T7"

gap

230430..230529
/estimated_length=unknown

ORIGIN

Alignment Scores:
Pred. No.: 1.71e+04 Length: 232670
Score: 43.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 89.6% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-9 (1-9) x AC129149 (1-232670)

Qy 1 GlyLeuProHisIleArgValPheLeu 9

|||||
Db 162875 GGGCTTCTCACTGAGACTTTCCTT 162849

RESULT 44

AC099417/c

LOCUS

AC099417 Rattus norvegicus clone CH230-2M18, WORKING DRAFT SEQUENCE.

DEFINITION

AC099417

VERSION

HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

1 (bases 1 to 252441)

Murphy,D.Marie., Metzker,M.Lee., Abranzon,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biewalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,B., Evans,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Gebregeorgis,E., Geer,K., Gill,R., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,C., Hamilton,J.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,A., Tingey,A., Trejos,Z., Umani,K.,
Valas,R., Vera,V., Villanana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

JOURNAL

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GABK
Center clone name: CH230-2M18
Summary Statistics
Assembly program: Atlas;
Consensus quality: 235495 bases at least Q40
Consensus quality: 236589 bases at least Q30

```

Consensus quality: 237381 bases at least Q20
Estimated insert size: 244427; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 252441: contig of 252441 bp in length.
*
FEATURES             source
    source
        1..252441
            /organism="Rattus norvegicus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10116"
            /clone="CH230-2M18"
        1..1396
            /note="wgs end_extension"
            clone_end:17"
        3241..3867
            /note="clone_boundary"
            clone_end:17"
            site:EcoRI
        end sequence:BH289046"
        248749..249386
            /note="clone_boundary"
            clone_end:Sp6
            site:EcoRI
        end sequence:BH289048"
        250556..252441
            /note="wgs end_extension"
            clone_end:Sp6"

ORIGIN

Alignment Scores:
Pred. No.:      1.86e+04      Length:      252441
Score:          43.00      Matches:      6
Percent Similarity: 100.0%      Conservative: 3
Best Local Similarity: 66.7%      Mismatches: 0
Query Match:      89.6%      Indels:      0
DB:              14      Gaps:      0

US-10-774-176-9 (1-9) x AC099417 (1-252441)

Qy 1 GlyLeuProHisIleArgValPheLeu 9
|||||:|||||:|||||:|||||:|||||:
Db 54623 GGGGNGCCCATGTGAGATCTTCTT 54597

RESULT 45
AC097908/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-150L5, *** SEQUENCING IN PROGRESS
** 10 unordered pieces.
ACCESSION AC097908
VERSION AC097908.7 GI:30520969
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 258569)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsebrook,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

```

```

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,W., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokeme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajz,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villaseana,D., Waldron,D., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 258569)
Worley,K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 258569)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24819718.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine

```

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GFON
Center clone name: CH230-150L5

----- Summary Statistics -----

Assembly program: Atlas 3.0;
Consensus quality: 238007 bases at least Q40
Consensus quality: 239775 bases at least Q30
Consensus quality: 240971 bases at least Q20
Estimated insert size: 246791; sum-of-confi-
guration coverage: 6x in Q20 bases; sum-of-con-

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 10 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

*	1	21332:	contig of 21332 bp in length
*	21833	21932:	gap of unknown length
*	21933	26406:	contig of 4474 bp in length
*	26407	26506:	gap of unknown length
*	26507	244662:	contig of 218156 bp in length
*	244663	244762:	gap of unknown length
*	244763	245801:	contig of 1039 bp in length
*	245802	245801:	gap of unknown length
*	245902	247216:	contig of 1315 bp in length
*	247217	247316:	gap of unknown length
*	247317	249218:	contig of 1902 bp in length
*	249219	249318:	gap of unknown length
*	249319	251049:	contig of 1731 bp in length
*	251050	251149:	gap of unknown length
*	251150	252468:	contig of 1319 bp in length
*	252469	252568:	gap of unknown length
*	252569	254150:	contig of 1592 bp in length
*	254151	254250:	gap of unknown length
*	254251	258569:	contig of 4319 bp in length.

FEATURES

```
1. 250503
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-15015"
```

misc feature

```

1.1755
/note="wgs_end_extension
clone_end:Sp6"
2220..2527
/note="clone_boundary
clone_end:Sp6
site:ECORI

```

misc feature

```

/note="wgs contig"
21833..21932
/estimated length=unknown

```

misc feature

```
/note="wgs_contig"  
26407. .26506  
/estimated length=unknown
```

misc_feature

```
/note="wgs_contig"  
216188. .216854  
/note="clone_boundary"
```

gap

gap	/estimated_length=unknown 245802. .245901
gap	/estimated_length=unknown 247217. .247316
gap	/estimated_length=unknown 249219. .249318
gap	/estimated_length=unknown 251050. .251149
gap	/estimated_length=unknown 252469. .252568
gap	/estimated_length=unknown 254151. .254250

ORIGIN

Alignment Scores:		
Pred. No.:	1.91e+04	258569
Score:	43.00	7
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	77.8%	Mismatches: 0
Query Match:	89.6%	Indels: 0
DB:	14	Gaps: 0

US-10-774-176-9 (1-9) x AC097908 (1-258569)

1 GlyLeuProHisIleArgValpHeLeu 9

db 204046 GGTGTACCCACCTAAGGGTCTTCTTA 204020

RESULT 46

AC122618/C

LOCUS

DEFINITION

unordered pieces.

AC122618

VERSION AC122618

KEYWORDS HTG; HTGS_PHASE1; HTGS_I

SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus

ORGANISM

Eukaryota; Metazoa
Mammalia: Eutheri

Mammalia; Sauria; Eucarchontoglires; Glires; Rodentia;
Sciurognathi: Muroidea: Muridae: Murinae: Rattus.

REFERENCE 1 (bases 1 to 263882)

AUTHORS

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Fleggg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, X., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulsegod, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathwney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwaekemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Perez, L., Pastermak, S., Paul, H., Perez, A., Perez, L., Prannkoch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodargbeyn, A., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Unpublished
2 (bases 1 to 263882)

Worley, K.C.

Direct Submission

Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 263882)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:24942501.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GRBP

Center clone name: CH230-65D16

Assembly program: Atlas 3.0;

Consensus quality: 256577 bases at least Q40

Consensus quality: 258619 bases at least Q30

Consensus quality: 260075 bases at least Q20

Estimated insert size: 266186; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

NOTE: This sequence may represent more than one clone.

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 258509: contig of 258509 bp in length
* 258609: gap of unknown length
* 258610 259774: contig of 1165 bp in length
* 259775 259874: gap of unknown length
* 259875 261421: contig of 1547 bp in length
* 261422 261521: gap of unknown length
* 261522 263882: contig of 2361 bp in length.

FEATURES

source

1. .263882

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-65D16"

258510..258609

/estimated_length=unknown

259775..259874

/estimated_length=unknown

261422..261521

/estimated_length=unknown

ORIGIN

Alignment Scores:

Pred. No.: 1.95e+04

Score: 43.00

Percent Similarity: 88.9%

Best Local Similarity: 88.9%

Query Match: 89.6%

DB: 14

Length: 263882

Matches: 8

Conservative: 0

Mismatches: 1

Indels: 0

Gaps: 0

US-10-774-176-9 (1-9) x AC122618 (1-263882)

QY 1 GlyLeuProHisIleArgValPheLeu 9

Db 90606 GGTCTTCCACATAGTGGCTTTCTC 90580

RESULT 47

AR520225

LOCUS

DEFINITION

Sequence 25185 from patent US 6703491.

AR520225

ACCESSION

AR520225.1

VERSION

GI:52455700

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 563)

AUTHORS

Homburger, S.A., Ebens, A.J. Jr., Erickson, C.S., Francis-Lang, H.L., Margolis, J.S., Reddy, B.P., Ruddy, D.A. and Buchman, A.R.

TITLE

Drosophila sequences

JOURNAL

Patent: US 6703491-A 25185 09-MAR-2004;

Exelixis, Inc.; South San Francisco, CA

FEATURES

Location/Qualifiers

source

1. .563

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 52.3

Score: 42.00

Percent Similarity: 100.0%

Best Local Similarity: 66.7%

Query Match: 87.5%

DB: 6

Length: 563

Matches: 6

Conservative: 3

Mismatches: 0

Indels: 0

Gaps: 0

US-10-774-176-9 (1-9) x AR520225 (1-563)

QY 1 GlyLeuProHisIleArgValPheLeu 9

Db 291 GGCCTTCTCCACCTACGAATATCTT 317

RESULT 48

AR504928

```

LOCUS       AR504928               1017 bp    DNA             linear      PAT 22-SEP-2004
DEFINITION   Sequence 9888 from patent US 6703491.
ACCESSION    AR504928
VERSION      AR504928.1  GI:52440403
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 1017)
AUTHORS      Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L.,
              Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.
TITLE        Drosophila sequences
JOURNAL      Patent: US 6703491-A 9888 09-MAR-2004;
              Exelisis, Inc.; South San Francisco, CA
FEATURES     source
              1..1017
               Location/Qualifiers
               /organism="unknown"
               /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      96.9      Length:      1017
Score:          42.00     Matches:      6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match:    87.5% Indels: 0
DB:             6 Gaps: 0

US-10-774-176-9 (1-9) x AR504928 (1-1017)
Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 745 GGCCTTCTCTCCTACCTACGAATATATCTT 771

RESULT 49
LOCUS       CQ600372               1704 bp    DNA             linear      PAT 02-FEB-2004
DEFINITION   Sequence 28130 from Patent WO0171042.
ACCESSION    CQ600372
VERSION      CQ600372.1  GI:41655561
KEYWORDS     Drosophila sp.
SOURCE       Drosophila sp.
ORGANISM     Drosophila sp.
REFERENCE    1
AUTHORS      Venter,J.C., Adams,M., Li,P.W. and Myers,E.W.
TITLE        Detection kits, such as nucleic acid arrays, for detecting the
              expression of 10,000 or more Drosophila genes and uses thereof
JOURNAL      Patent: WO 0171042-A 28130 27-SEP-2001;
              PE Corporation (NY) (US)
FEATURES     source
              1..1704
               Location/Qualifiers
               /organism="Drosophila sp."
               /mol_type="unassigned DNA"
               /db_xref="taxon:7242"

ORIGIN
Alignment Scores:
Pred. No.:      166      Length:      1704
Score:          42.00     Matches:      6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match:    87.5% Indels: 0
DB:             6 Gaps: 0

US-10-774-176-9 (1-9) x CQ600372 (1-1704)
Qy 1 GlyLeuProHisIleArgValPheLeu 9
Db 720 GGCCTTCTCTCCTACCTACGAATATATCTT 746

```

```

RESULT 50
LOCUS       BT006007               1980 bp    mRNA            linear      INV 28-MAR-2003
DEFINITION   Drosophila melanogaster RE18708 full insert cDNA.
ACCESSION    BT006007
VERSION      BT006007.1  GI:293335982
KEYWORDS     FLY CDNA.
SOURCE       Drosophila melanogaster (fruit fly)
ORGANISM     Drosophila melanogaster
REFERENCE    1 (bases 1 to 1980)
AUTHORS      Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
              Champe,M., Chavez,C., Dorsett,V., Dreanek,D., Farfan,D., Friso,E.,
              George,R., Gonzalez,M., Guarin,H., Krommiller,B., Li,P., Liao,G.,
              Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
              Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
              and Celniker,S.
FEATURES     Direct Submission
              Submitted (28-MAR-2003) Berkeley Drosophila Genome Project,
              Lawrence Berkeley National Laboratory, One Cyclotron Road,
              Berkeley, CA 94720, USA
              Sequence submitted by:
              Berkeley Drosophila Genome Project
              Lawrence Berkeley National Laboratory
              Berkeley, CA 94720
              This clone was sequenced as part of a high-throughput process to
              sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
              Science 2000). The sequence has been subjected to integrity checks
              for sequence accuracy, presence of a polyA tail and contiguity
              within 100 kb in the genome. Thus we believe the sequence to
              reflect accurately this particular cDNA clone. However, there are
              artifacts associated with the generation of cDNA clones that may
              have not been detected in our initial analyses such as internal
              priming, priming from contaminating genomic DNA, retained introns
              due to reverse transcription of unspliced precursor RNAs, and
              reverse transcriptase errors that result in single base changes.
              For further information about this sequence, including its location
              and relationship to other sequences, please visit our Web site
              (http://fruitfly.berkeley.edu) or send email to
              cdna@fruitfly.berkeley.edu.
              Location/Qualifiers
              1..1980
               /organism="Drosophila melanogaster"
               /mol_type="mRNA"
               /strain="y; cn bw sp"
               /db_xref="taxon:7227"
              1..1980
               /gene="Ugt86Da"
               /db_xref="FLYBASE:FBgn0040259"
               142..1728
               /gene="Ugt86Da"
               /note="Longest ORF"
               /codon_start=1
               /product="RE18708p"
               /protein_id="AAO74690.1"
               /db_xref="GI:293335983"
               /db_xref="FLYBASE:FBgn0040259"
               /translation="MKPTAGQTGFLALLFCLLSCVSAYNYLVLVLTARSHYVGSAL
               LAKGLAAGHQVTIISPPELAKPKINIKDVPAKSLTSMQGRIANLQSSKEPIIKQI
               INFHEWGEITELLKPSVIELMKSNOTEDAVISEVFLNEAHFGFAHFKAFLIGL
               TFGALSWNTDLVGSPPSPYVPSALLKFSDRMSLVERVQNOALFYEVFLNRYLPR
               QEVLYKFPNNKQDFYDMRKNTALVLLNQHVLSLFFRPYSNMIEVGGMHNRRKQP
               LPKDILEFIEAGHGVIYFSMGSNLKSITLPLEKQALIDTFAQLKQVLMWPKEDFL
               PKGANVFISDFPQDDILAHDNVLAFTHGGLLSITTHYRKPFGVPIPFQDQFLN
               MARAEQNGYGVTVHYEELSSAKLLAAIKIINNPEATQVRDMSDRYRQDQQTFLERA
               VVWEVSHRHKGAKYLRASQDLNFIQHNLDLMLLYGLIIFVLYCYIFLLRLVRL
               LQELFIKDTSKPKQAKGN"

FEATURES     source
              1..1980
               /organism="Drosophila melanogaster"
               /mol_type="mRNA"
               /strain="y; cn bw sp"
               /db_xref="taxon:7227"
              1..1980
               /gene="Ugt86Da"
               /db_xref="FLYBASE:FBgn0040259"
               142..1728
               /gene="Ugt86Da"
               /note="Longest ORF"
               /codon_start=1
               /product="RE18708p"
               /protein_id="AAO74690.1"
               /db_xref="GI:293335983"
               /db_xref="FLYBASE:FBgn0040259"
               /translation="MKPTAGQTGFLALLFCLLSCVSAYNYLVLVLTARSHYVGSAL
               LAKGLAAGHQVTIISPPELAKPKINIKDVPAKSLTSMQGRIANLQSSKEPIIKQI
               INFHEWGEITELLKPSVIELMKSNOTEDAVISEVFLNEAHFGFAHFKAFLIGL
               TFGALSWNTDLVGSPPSPYVPSALLKFSDRMSLVERVQNOALFYEVFLNRYLPR
               QEVLYKFPNNKQDFYDMRKNTALVLLNQHVLSLFFRPYSNMIEVGGMHNRRKQP
               LPKDILEFIEAGHGVIYFSMGSNLKSITLPLEKQALIDTFAQLKQVLMWPKEDFL
               PKGANVFISDFPQDDILAHDNVLAFTHGGLLSITTHYRKPFGVPIPFQDQFLN
               MARAEQNGYGVTVHYEELSSAKLLAAIKIINNPEATQVRDMSDRYRQDQQTFLERA
               VVWEVSHRHKGAKYLRASQDLNFIQHNLDLMLLYGLIIFVLYCYIFLLRLVRL
               LQELFIKDTSKPKQAKGN"

ORIGIN
Alignment Scores:
Pred. No.:      194      Length:      1980

```

Score: 42.00 Matches: 6
Percent Similarity: 100.0% Conservative: 3
Best Local Similarity: 66.7% Mismatches: 0
Query Match: 87.5% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-9 (1-9) x BT006007 (1-1980)

QY 1 GlyLeuProHisIleArgValPheLeu 9
Db 744 GGCCTTCCTCACCTACGAAATATCTT 770

Search completed: April 25, 2006, 20:44:04
Job time : 3134.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds
(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-8

Perfect score: 41

Sequence: 1 AIFLLVLVL 9

Scoring table: BIOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abss/ABSSWEB spo01/US10774176/runat_24042006.165112.19185/app_query.fasta_1
-DB=N Geneseq -QEXT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNIT8=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05p
-USER=US10774176 @CGEN 1.1 3463 @runat_24042006.165112.19185 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOXT=6 -DELEXT=7

Database :

N Geneseq 21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	246	10 ADK11641	Adk11641 Breast ca
2	41	100.0	475	13 ADU11677	Adu11677 Solid tum
3	41	100.0	901	3 AAA27060	Aaa27060 Canine 5T
4	41	100.0	927	6 ABT07721	Abt07721 Breast ca

5	41	100.0	927	8 ABX76333	Abx76333 Lung canc
6	41	100.0	927	10 ADB80503	Adb80503 Ovarian c
7	41	100.0	927	11 ADN38723	Adn38723 Cancer/an
8	41	100.0	973	8 AAD56198	Aad56198 Human LRR
9	41	100.0	1156	6 ABV99349	Abv99349 Human NOV
10	41	100.0	1260	6 ABK87175	Abk87175 cDNA enco
11	41	100.0	1260	10 ADB97513	Adb97513 Feline 5T
12	41	100.0	1260	10 ADB97452	Adb97452 DNA enco
13	41	100.0	1263	3 AAA27058	Aaa27058 Human 5T4
14	41	100.0	1263	4 AAF89736	Aaf89736 Nucleotid
15	41	100.0	1263	6 ABK87174	Abk87174 cDNA enco
16	41	100.0	1281	3 AAA27059	Aaa27059 Mouse 5T4
17	41	100.0	1331	8 AAD56199	Aad56199 Human LRR
18	41	100.0	2020	10 ADJ56299	Adj56299 Human CDN
19	41	100.0	2053	8 ACC51052	Acc51052 Human bla
20	41	100.0	2053	8 ABX76332	Abx76332 Lung canc
21	41	100.0	2053	8 AAD56197	Aad56197 Human LRR
22	41	100.0	2053	8 AAD56200	Aad56200 Human LRR
23	41	100.0	2053	11 ADN38721	Adn38721 Cancer/an
24	41	100.0	2053	12 ADL06473	Adl06473 Human tum
25	41	100.0	2053	12 ADN03961	Adn03961 Antipeori
26	41	100.0	2053	13 ADR25444	Adr25444 Breast ca
27	41	100.0	2053	13 ACN38510	Acn38510 Tumour-as
28	41	100.0	2053	13 ADV35098	Adv35098 Human CDN
29	41	100.0	2338	5 AAS87175	Aas87175 DNA enco
30	41	100.0	2359	4 AAK94253	Aak94253 Human ful
31	41	100.0	2359	12 ADL30831	Adl30831 Full leng
32	41	100.0	2361	4 AAK94254	Aak94254 Human ful
33	41	100.0	2361	12 ADI26162	Adi26162 Human CDN
34	41	100.0	2361	12 ADL30833	Adl30833 Full leng
35	41	100.0	2557	12 ADI26160	Adi26160 Human CDN
36	41	100.0	2557	12 ADI26158	Adi26158 Human CDN
37	40	97.6	19142	2 AAX20580	Aax20580 Polynucle
38	37	90.2	204	10 ADH84583	Adh84583 Enterococ
39	37	90.2	1058	4 ABL06215	AbL06215 Drosophil
40	37	90.2	1421	3 AAC46422	Aac46422 Arabidops
41	37	90.2	1422	3 AAC36467	Aac36467 Arabidops
42	37	90.2	1689	13 ADT15316	Adt15316 Plant CDN
43	37	90.2	3298	4 ABL06214	AbL06214 Drosophil
44	37	90.2	6224	6 ABL33308	AbL33308 Human imm
45	37	90.2	6224	6 ABL54355	AbL54355 Chemicall
46	37	90.2	32768	2 AAX13336	Aax13336 Enterococ
47	37	90.2	32768	6 ABS99131	AbS99131 Enterococ
48	37	90.2	92139	6 AAD31364	Aad31364 92Kb gene
49	37	90.2	110000	11 ACN44006	Acn44006 Human gen
50	37	90.2	127369	11 ADT42063_16	Adt42063_16 o
51	37	90.2	130320	10 ADF11613	Adf11613 Human scl
52	36	87.8	582	6 ABK77524	Abk77524 Bacillus
53	36	87.8	667	10 ABZ83591	Abz83591 Toxicolog
54	36	87.8	1018	3 AAA38554	Aaa38554 Actinobac
55	36	87.8	1065	13 ADS59278	Ads59278 Bacterial
56	36	87.8	1260	12 ADO07184	Ado07184 Fusarium
57	36	87.8	1335	8 ACA28617	Aca28617 Prokaryot
58	36	87.8	1395	2 AAQ66812	Aaq66812 AmSPV sph
59	36	87.8	1395	6 ABL56230	AbL56230 AmSPV cor
60	36	87.8	1395	6 ABT08013	AbT08013 DNA enco
61	36	87.8	2673	8 ACA29127	Aca29127 Prokaryot
62	36	87.8	2679	13 ADT05260	Adt05260 Haemophil
63	36	87.8	2730	10 ADB80434	Adb80434 Rat CLCA1
64	36	87.8	2730	12 ADI29518	Adi29518 Rat CLCA1
65	36	87.8	2730	14 AEA35975	Aea35975 Novel tra
66	36	87.8	3267	12 ADK52131	Adk52131 Mouse ato
67	36	87.8	5270	4 AAK64901	Aak64901 Human imm
68	36	87.8	6766	2 AAV14507	Aav14507 AmSPV ent
69	36	87.8	6768	2 AAQ28301	Aaq28301 AmSPV Sph
70	36	87.8	6768	2 AAV14517	Aav14517 AmSPV ent
71	36	87.8	8267	4 ABL21660	AbL21660 Drosophil
72	36	87.8	8455	2 AAZ10081	Aaz10081 Amsacta m
73	36	87.8	8457	6 AAQ66797	Aaq66797 AmSPV sph
74	36	87.8	8457	6 ABT07998	Abt07998 DNA of th
75	36	87.8	32460	12 ADQ07738	Adq07738 Nucleotid
76	36	87.8	45736	13 ABD33564	Abd33564 Murine ca
77	36	87.8	50000	6 ABL56202	AbL56202 AmSPV gen

78	36	87.8	76500	12	ADP45592	Adp45592 Human MAP	151	35	85.4	4435	2	AAQ29686	AaQ29686 CCV-C54 s
79	36	87.8	110050	2	AAT42063_10	Continuation (11 o	152	35	85.4	4476	4	AAH73361	AaH73361 Human cer
80	36	87.8	110950	14	ADX98569	Adx98569 Human mit	153	35	85.4	4486	2	AAQ35118	AaQ35118 Encodes E
C 81	36	87.8	125439	6	AQO88177	Aqc88177 Human oet	154	35	85.4	4500	2	AAQ29684	AaQ29684 CCV-6 epi
82	36	87.8	191996	13	ADT05647	Adt05647 Haemophil	155	35	85.4	4673	14	ADZ64440	Adz64440 Human can
83	36	87.8	200418	11	ACN44226	Acn44226 Human gen	156	35	85.4	5035	11	ADL22612	Adl22612 Human dis
84	36	87.8	245531	13	ABD33022	Abd33022 Human can	157	35	85.4	5035	14	ADY18842	Ady18842 DNA encod
85	35	85.4	263	6	ABV96123	Abv96123 Human pan	158	35	85.4	5529	14	ADZ64439	Adz64439 Human can
86	35	85.4	266	6	ABL37787	Ab137787 Human col	159	35	85.4	5554	8	ABV73333	Abv73333 M. sexta
C 87	35	85.4	266	6	ABL37964	Ab137964 Human col	160	35	85.4	5873	14	ADZ64433	Adz64433 Human can
C 88	35	85.4	286	10	ADF85681	Adf85681 Human ade	161	35	85.4	5967	14	ADZ64437	Adz64437 Human can
89	35	85.4	356	14	ADV74945	Adv74945 Human col	162	35	85.4	5967	14	ADZ64436	Adz64436 Human can
C 90	35	85.4	411	11	ABD09155	Abd09155 Pseudomon	163	35	85.4	6021	14	ADZ64432	Adz64432 Human can
C 91	35	85.4	412	10	ADP80089	Adp80089 Leukemia	164	35	85.4	6084	14	ADZ64434	Adz64434 Human can
C 92	35	85.4	459	5	AAF68129	Aaf68129 Human lun	165	35	85.4	6086	14	ADZ64438	Adz64438 Human can
C 93	35	85.4	459	6	ABK38040	Abk38040 CDNA enco	166	35	85.4	6590	2	AAI12969	Aai12969 Enterococ
C 94	35	85.4	459	8	ACA10369	Act10369 Human lun	167	35	85.4	6590	6	ABN98764	Abn98764 Enterococ
C 95	35	85.4	459	8	ACH99320	Ach99320 Lung canc	168	35	85.4	6641	6	ABN80002	Abn80002 Human che
C 96	35	85.4	459	10	ADH45566	Adh45566 Human lun	169	35	85.4	6730	2	AAV74362	Aav74362 Staphyloc
C 97	35	85.4	459	12	ADH72103	Adh72103 Human lun	C 170	35	85.4	7119	13	ADT04702	Adt04702 House mou
C 98	35	85.4	459	13	ADJ19485	Adj19485 Human lun	C 171	35	85.4	9341	13	ADT05504	Adt05504 Haemophil
C 99	35	85.4	472	9	ACH34512	Ach34512 Human end	C 172	35	85.4	9434	14	ADZ64430	Adz64430 Human can
100	35	85.4	475	9	ACH19702	Ach19702 Human adu	C 173	35	85.4	9471	2	AAV57356	Aav57356 FIV-NC5U1
C 101	35	85.4	498	3	AAC94753	Aac94753 Cat flea	C 174	35	85.4	9474	2	AAV57355	Aav57355 FIV-NC5U1
C 102	35	85.4	498	12	ADL11424	Adl11424 Cat flea	C 175	35	85.4	10688	4	ABL03406	Ab103406 Drosophil
C 103	35	85.4	523	6	ABV87224	Abv87224 Human col	C 176	35	85.4	12135	4	ABL03408	Ab103408 Drosophil
C 104	35	85.4	523	6	ABW86969	Abw86969 Human col	C 177	35	85.4	12755	4	ABL09786	Ab109786 Drosophil
C 105	35	85.4	531	6	ABT10425	Abt10425 Human bre	C 178	35	85.4	17384	4	ABL04140	Ab104140 Drosophil
C 106	35	85.4	531	13	ADT90105	Adt90105 Human gen	179	35	85.4	31449	4	ABL09172	Ab109172 Drosophil
C 107	35	85.4	531	13	ADT90207	Adt90207 Human gen	180	35	85.4	59725	11	ACN44390	Acn44390 Human gen
C 108	35	85.4	531	13	ADT90253	Adt90253 Human gen	C 181	35	85.4	109565	13	ABD33086	Abd33086 Murine ca
C 109	35	85.4	583	4	AH343300	Aah343300 Human col	C 182	35	85.4	110000	2	AAT42063_07	Continuation (8 of
C 110	35	85.4	627	8	ACA30567	Act30567 Prokaryot	C 183	35	85.4	110000	2	AAZ20248_04	Continuation (5 of
111	35	85.4	642	8	ACA50465	Act50465 Prokaryot	C 184	35	85.4	110000	11	ACN44934_2	Continuation (3 of
112	35	85.4	720	6	ABN67450	Abn67450 Streptoco	C 185	35	85.4	110000	11	ACN44014_2	Continuation (3 of
113	35	85.4	729	6	ABK55412	Abk55412 Human col	186	35	85.4	110000	12	ADQ97050_4	Continuation (5 of
C 114	35	85.4	731	8	ACA29429	Act29429 Prokaryot	187	35	85.4	110000	2	ADZ46976_4	Continuation (5 of
C 115	35	85.4	913	5	AH81625	Aah81625 Human dif	188	35	85.4	112777	2	AAZ20249	AAZ20249 Borrelia
C 116	35	85.4	1020	4	ABL20357	Ab120357 Drosophil	189	35	85.4	135356	13	ADT05646	Adt05646 Haemophil
C 117	35	85.4	1068	8	ACF74770	Act74770 Staphyloc	C 190	35	85.4	158980	13	ADZ44537	Adz44537 Human bre
118	35	85.4	1104	11	ABD09121	Abd09121 Pseudomon	191	34	82.9	57	2	AAQ27380	AaQ27380 17-alpha-
119	35	85.4	1118	6	ABL90734	Ab190734 Human pol	192	34	82.9	57	2	AAQ55118	AaQ55118 Human liv
120	35	85.4	1126	3	ABC36606	Abc36606 Arabidops	C 193	34	82.9	245	3	AAAS7025	AAAS7025 Human col
121	35	85.4	1147	4	AAI26726	Aai26726 Human pol	C 194	34	82.9	245	6	ABT12447	Abt12447 Orestes s
122	35	85.4	1149	6	ABA90338	AbA90338 Human bre	C 195	34	82.9	245	10	ACD91741	Act91741 Human col
C 123	35	85.4	1362	8	ACA29167	Act29167 Prokaryot	C 196	34	82.9	255	5	AAH81911	Aah81911 Rat diffe
C 124	35	85.4	1478	13	ADM50653	Adm50653 Human vac	197	34	82.9	258	3	AAA41905	Aaa41905 Human sec
C 125	35	85.4	1617	13	ADR73382	Adr73382 Thale cre	198	34	82.9	259	5	ABV15390	Abv15390 Human pro
C 126	35	85.4	1743	3	ACF78147	Act78147 Human can	C 199	34	82.9	272	6	ABL82747	Ab182747 Human ova
C 127	35	85.4	1985	13	ACN40031	Actn40031 Tumour-as	200	34	82.9	281	5	ABV06221	Abv06221 Human pro
C 128	35	85.4	2000	11	ACL37602	Actl37602 Rice stre	201	34	82.9	293	6	ABL73178	Ab173178 Corn tass
C 129	35	85.4	2018	6	ABZ57565	Abz57565 Human vac	202	34	82.9	324	2	AAV88708	Aav88708 EST clone
130	35	85.4	2094	11	ABD08971	Abd08971 Pseudomon	C 203	34	82.9	365	6	ABL66653	Ab166653 Lung canc
C 131	35	85.4	2100	11	ADN95832	Adn95832 Human BEC	204	34	82.9	385	5	ACH50384	Act50384 Human leu
C 132	35	85.4	2120	10	ADC14280	Act14280 Human enz	205	34	82.9	387	5	AAF66672	Aaf66672 Novel hum
C 133	35	85.4	2141	14	ADZ64445	Adz64445 Human can	206	34	82.9	399	5	AAH81658	Aah81658 Human dif
C 134	35	85.4	2447	12	ADQ15049	Adq15049 Human can	207	34	82.9	420	8	ABZ20591	Abz20591 A thalian
C 135	35	85.4	2447	13	ADR24675	Adr24675 Breast ca	C 208	34	82.9	421	9	ACH46125	Act46125 Human inf
C 136	35	85.4	2447	13	ADP23198	Adp23198 PRO polyp	209	34	82.9	436	13	ADR13078	Adr13078 Human can
137	35	85.4	2670	14	ADZ64444	Adz64444 Human can	210	34	82.9	442	9	ACH27369	Act27369 Human adu
138	35	85.4	2693	12	ADH43911	Adh43911 Chicken D	211	34	82.9	449	5	ABV45237	Abv45237 Human pro
139	35	85.4	2706	4	AA826198	Aa826198 Human cdn	212	34	82.9	449	5	ABV45237	Abv45237 Human pro
140	35	85.4	2706	8	ABX73539	Abx73539 Human nov	213	34	82.9	471	13	ADR13072	Adr13072 Human can
C 141	35	85.4	2725	5	AA833254	Aa833254 DNA encod	C 214	34	82.9	485	9	ACH35498	Act35498 Human end
C 142	35	85.4	3020	4	ABE120356	Ab120356 Drosophil	215	34	82.9	496	14	ABE66278	Aeb66278 Rice geno
143	35	85.4	3356	4	AAK65287	Aak65287 Human hum	216	34	82.9	503	12	ACH78264	Act78264 Human gen
144	35	85.4	3993	13	ADU01816	Adu01816 Novel hum	217	34	82.9	521	14	ACL62829	Act62829 Human col
145	35	85.4	4116	14	ADP64441	Adp64441 Human can	218	34	82.9	556	4	AAH11065	Aah11065 Human cdn
146	35	85.4	4195	12	ADP88212	Adp88212 Hansenula	219	34	82.9	561	11	ACN85824	Actn85824 Breast ca
147	35	85.4	4195	13	ADU00243	Adu00243 Gamma-glu	C 220	34	82.9	587	5	ABV53188	Abv53188 Human pro
148	35	85.4	4359	2	AAQ52447	Aaq52447 Canine co	C 221	34	82.9	623	6	ABT10957	Abt10957 Human bre
149	35	85.4	4359	2	AAQ50617	Aaq50617 Canine co	222	34	82.9	624	11	ACL30850	Actl30850 Rice abio
150	35	85.4	4363	12	ADK42718	Adk42718 Canine co	223	34	82.9	711	4	AA503150	Aas03150 Human EPO

224	34	82.9	774	9	ADA311208	Ada31208 DNA encod	C 297	34	82.9	3426	10	ADE72601	Ad72601 Human end
225	34	82.9	788	13	ADx10420	Adx10420 Plant full	C 298	34	82.9	3440	12	AD116295	Ad116295 Human nuc
C 226	34	82.9	792	8	ACA28691	Aca28691 Prokaryot	C 299	34	82.9	3477	6	ABL70013	Ab170013 Pancreas
C 227	34	82.9	810	10	ADH84155	Adh84155 Enterococ	C 300	34	82.9	3477	13	ABK84207	Abk84207 Human cdn
228	34	82.9	826	8	AB252298	Ab252298 Aspergill	301	34	82.9	3477	13	ADRO7931	Adr07931 Full leng
229	34	82.9	831	5	ADH81310	Aah81310 Escherich	302	34	82.9	3486	12	AQ96439	Adq96439 T cell ac
C 230	34	82.9	833	8	ACA27609	Aca27609 Prokaryot	303	34	82.9	3496	5	AA888536	Aas888536 DNA encod
231	34	82.9	962	3	AA54674	Aac54674 Arabidops	304	34	82.9	3616	13	AD547830	Ad547830 Bacterial
232	34	82.9	964	3	AA53487	Aac53487 Arabidops	C 305	34	82.9	3621	4	ABL08516	Ab108516 Drosophil
233	34	82.9	981	6	ABK35066	Abk35066 Human cdn	C 306	34	82.9	3624	12	ADQ25456	Adq25456 Human sof
234	34	82.9	1002	3	AAF14627	Aaf14627 Aspergill	C 307	34	82.9	3636	4	ABL14812	Ab114812 Drosophil
235	34	82.9	1002	13	ADU58668	Adu58668 Aspergill	C 308	34	82.9	3637	1	AA911475	Aan91475 Human top
236	34	82.9	1002	14	AD296671	Ad296671 Aspergill	C 309	34	82.9	3637	1	AA911475	Aan91475 Human top
237	34	82.9	1024	10	AD283341	Ab283341 Toxicolog	C 310	34	82.9	3645	2	AA911432	Aaq91432 Variant h
C 238	34	82.9	1059	9	ADA311134	Ada311134 DNA encod	C 311	34	82.9	3645	2	AA959395	Aax59395 Human top
239	34	82.9	1158	10	ACF70521	Act70521 Photorhab	C 312	34	82.9	3734	12	ADL82892	Adl82892 Human PRO
240	34	82.9	1258	13	ADT17845	Adt17845 Plant cdn	C 313	34	82.9	3734	13	ADP56067	Adp56067 Human PRO
241	34	82.9	1508	6	ABK99437	Abk99437 Human CYP	C 314	34	82.9	3739	8	ABZ23989	Abz23989 Human top
242	34	82.9	1596	11	ACL28872	Acl28872 Rice abio	C 315	34	82.9	3740	8	ACF62748	Acf62748 Cancer ba
243	34	82.9	1626	6	AB232545	Ab232545 Candida a	C 316	34	82.9	3740	9	ADA02679	Ada02679 Human TOP
C 244	34	82.9	1683	13	ADX60580	Adx60580 Plant ful	C 317	34	82.9	3740	10	ADB87956	Adb87956 Human UGT
245	34	82.9	1793	8	ABX34541	Abx34541 Human mdd	C 318	34	82.9	3740	10	ADB72417	Adb72417 Human TIO
246	34	82.9	1845	9	ADA31248	Ada31248 DNA encod	C 319	34	82.9	3740	10	ADB96939	Adb96939 Human MDR
247	34	82.9	1859	6	ABQ33833	Abq33833 Oligonucl	C 320	34	82.9	3740	10	ADB92130	Adb92130 Human MDR
C 248	34	82.9	1859	6	ABQ33832	Abq33832 Oligonucl	C 321	34	82.9	3740	10	ADB95927	Adb95927 Human DNA
249	34	82.9	1969	2	AAV34250	Aav34250 Human sec	C 322	34	82.9	3756	10	AD102714	Ad102714 Human cdn
250	34	82.9	1969	8	ACD08121	Acc08121 cDNA enco	C 323	34	82.9	3831	10	AD72593	Ad72593 Human end
251	34	82.9	1969	14	AD121307	Ad121307 Human sec	C 324	34	82.9	4114	10	ADC37627	Adc37627 Human nuc
252	34	82.9	2000	6	AB214952	Ab214952 Arabidops	C 325	34	82.9	4515	13	ADT47715	Adt47715 Bacterial
253	34	82.9	2000	8	ADA72207	Ada72207 Rice gene	C 326	34	82.9	4568	4	ABL07472	Ab107472 Drosophil
C 254	34	82.9	2000	8	ADA72859	Ada72859 Rice gene	C 327	34	82.9	4576	10	ADC30036	Adc30036 Human nov
C 255	34	82.9	2000	12	ADJ41229	Adj41229 Plant cdn	C 328	34	82.9	4611	10	ADB72588	Adb72588 Human end
C 256	34	82.9	2006	3	AA81092	Aac81092 Human sec	C 329	34	82.9	4696	2	AAAX03041	Aax03041 Human IL-
257	34	82.9	2111	2	AAV69308	Aav69308 Human EPR	C 330	34	82.9	4733	6	ABX92020	Abx92020 Lung spec
258	34	82.9	2111	4	AA503149	Aas03149 Human EPO	331	34	82.9	4901	4	ABL05477	Ab105477 Drosophil
C 259	34	82.9	2169	13	ADX64771	Adx64771 Plant ful	332	34	82.9	4901	10	ADK11343	Adk11343 Drosophil
C 260	34	82.9	2178	8	ACA28141	Aca28141 Prokaryot	C 333	34	82.9	4935	4	ABL15050	Ab115050 Drosophil
261	34	82.9	2200	11	AD122572	Ad122572 Human dis	334	34	82.9	4992	8	ABZ20576	Abz20576 A thalian
C 262	34	82.9	2205	5	ABV23980	Abv23980 Human pro	C 335	34	82.9	5357	4	ABL18826	Ab118826 Drosophil
C 263	34	82.9	2205	6	AAV29865	Abv29865 Human pro	C 336	34	82.9	5504	6	ABL170349	Ab1170349 Chemical
C 264	34	82.9	2298	9	ADA02680	Ada02680 Human TOP	C 337	34	82.9	5504	6	AA561308	Aas61308 Human gen
C 265	34	82.9	2298	10	ADB72418	Adb72418 Human TIO	C 338	34	82.9	5825	6	ABK28283	Abk28283 DNA trans
C 266	34	82.9	2298	10	AD95928	Ad95928 Human TOP	C 339	34	82.9	6026	6	ABL02768	Ab102768 Drosophil
267	34	82.9	2337	2	AAQ11827	Aaq11827 Fission y	340	34	82.9	6029	6	ABL92256	Ab192256 Chemical
268	34	82.9	2342	2	AAV69309	Aav69309 Human EPR	341	34	82.9	6029	6	AD22325	Ad22325 Chemical
269	34	82.9	2342	2	AAV69307	Aav69307 Human EPR	342	34	82.9	6145	6	ABL32972	Ab132972 Human imm
270	34	82.9	2342	4	AA503148	Aas03148 Human EPO	C 343	34	82.9	6650	6	ABK31413	Abk31413 Signal tr
271	34	82.9	2342	10	ABZ22649	Abz22649 Human EPO	344	34	82.9	7287	2	AAAX13147	Aax13147 Enterococ
272	34	82.9	2378	2	AAV34188	Aav34188 Human sec	345	34	82.9	7287	6	ABX98942	Abx98942 Enterococ
273	34	82.9	2378	8	ACD08059	Acc08059 cDNA enco	346	34	82.9	7459	6	ABK31383	Abk31383 Signal tr
274	34	82.9	2378	14	AD12245	Ad12245 Human sec	347	34	82.9	7814	4	AA546530	Aas46530 Tumour su
C 275	34	82.9	2390	4	ABL02769	Ab102769 Drosophil	348	34	82.9	8256	4	ABL05476	Ab105476 Drosophil
C 276	34	82.9	2400	2	AA559592	Aax55952 DNA encod	349	34	82.9	8919	4	ABL20699	Ab120699 Drosophil
277	34	82.9	2405	2	AAV58763	Aav58763 Human sec	350	34	82.9	8968	4	ABL13297	Ab113297 Drosophil
C 278	34	82.9	2467	10	ADB62316	Adb62316 Human cdn	351	34	82.9	9417	2	AAAX36849	Aax36849 Human XLI
C 279	34	82.9	2469	6	ABK84206	Abk84206 Human cdn	352	34	82.9	9417	8	ABX76234	Abx76234 Lung canc
C 280	34	82.9	2590	12	ADL61781	Adl61781 cDNA enco	353	34	82.9	9417	14	ADV70098	Adv70098 Tumor-ass
281	34	82.9	2595	11	ADI30928	Adi30928 Human cdn	C 354	34	82.9	9443	4	AAK68414	Aak68414 Human imm
C 282	34	82.9	2595	13	ADS82995	Ad82995 Human cdn	C 355	34	82.9	9445	4	AAK68415	Aak68415 Human imm
C 283	34	82.9	2870	4	AAI99289	Aai99289 Human exc	356	34	82.9	9829	3	AAZ35271	Aaz35271 Soybean r
C 284	34	82.9	2870	5	AAI63639	Aai63639 Human kid	357	34	82.9	10433	6	ABL32379	Ab132379 Human imm
285	34	82.9	2878	3	AAZ93078	Aaz93078 Partial s	C 358	34	82.9	10703	4	AA546260	Aas46260 Drosophil
C 286	34	82.9	2909	10	ADE72597	Ad72597 Human end	359	34	82.9	12144	4	AA546260	Aas46260 DNA encod
C 287	34	82.9	2926	4	ABL21192	Ab121192 Drosophil	360	34	82.9	12192	4	AA528195	Aas28195 Genomic B
C 288	34	82.9	3024	10	ADE72595	Ad72595 Human end	361	34	82.9	12192	10	ADG41391	Adg41391 Human res
C 289	34	82.9	3048	10	ADE72599	Ad72599 Human end	362	34	82.9	12192	11	AD197165	Ad197165 Human res
290	34	82.9	3088	14	ADY18753	Ady18753 DNA encod	363	34	82.9	12870	6	ABK39984	Abk39984 Human che
C 291	34	82.9	3106	10	AD72594	Ad72594 Human end	364	34	82.9	12870	6	ABL70230	Ab170230 Chemical
C 292	34	82.9	3130	10	ADE72584	Ad72584 Human end	365	34	82.9	13317	4	AA546682	Aas46682 Tumour su
C 293	34	82.9	3297	4	AA94453	Aa94453 Human hyd	366	34	82.9	13317	6	AA561367	Aas61367 Human gen
C 294	34	82.9	3408	10	ADE72602	Ad72602 Human end	C 367	34	82.9	14708	4	ABL13296	Ab113296 Drosophil
C 295	34	82.9	3412	10	ADE72598	Ad72598 Human end	C 368	34	82.9	14872	2	AAV52205	Aav52205 Streptoco
296	34	82.9	3424	4	AAK85925	Aak85925 Human imm	369	34	82.9	17721	6	ABL33729	Ab133729 Human imm

c 370	34	82.9	18683	6	ABL32313	Abi32313 Human imm	443	33	80.5	332	6	ABL84463	Abi84463 Human ova
c 371	34	82.9	18683	6	ABL54334	Abi54334 Chemical	444	33	80.5	365	5	AAH94122	Aah94122 Human foe
c 372	34	82.9	20001	14	ABE96525	Abe96525 Human IFI	445	33	80.5	383	13	ACF85051	Acf85051 Human STR
c 373	34	82.9	20486	4	ABL20698	Abi20698 Drosophil	c 446	33	80.5	388	4	AAI17471	Aai17471 Human bre
c 374	34	82.9	21354	4	AA846815	Aa846815 Tumour su	447	33	80.5	389	4	AAI91159	Aai91159 Human pol
c 375	34	82.9	24939	6	ABL70570	Abi70570 Chemical	448	33	80.5	398	4	AAI91177	Aai91177 Human pol
c 376	34	82.9	25559	10	ACF65388_12	Continuation (13 o	449	33	80.5	424	13	ACF83784	Acf83784 Human STR
c 377	34	82.9	26554	14	ABE33408	Aeb33408 Human gen	450	33	80.5	427	3	AAAC03357	Aac03357 Human sec
c 378	34	82.9	30515	4	ABK42321	Abk42321 Genomic s	c 451	33	80.5	429	2	AAAX40906	Aax40906 Human sec
c 379	34	82.9	30515	4	AAK68732	Aak68732 Human imm	c 452	33	80.5	429	5	ABV49095	Abv49095 Human pro
c 380	34	82.9	30515	4	AAK85029	Aak85029 Human imm	453	33	80.5	431	13	ADX35762	Adx35762 Plant ful
c 381	34	82.9	30515	9	ADB60477	Adb60477 Connectiv	454	33	80.5	451	6	ABL36474	Abi36474 Human col
c 382	34	82.9	32768	2	AAI13060	Aai13060 Enterococ	c 455	33	80.5	456	9	ACH33820	Ach33820 Human end
c 383	34	82.9	32768	6	AB898855	Ab898855 Enterococ	c 456	33	80.5	473	3	AAI93634	Aai93634 Human pol
c 384	34	82.9	34688	6	ABQ67059	Abq67059 Human ang	c 457	33	80.5	473	3	AAAC52676	Aac52676 Arabidops
c 385	34	82.9	35678	4	AAK84996	Aak84996 Human imm	c 458	33	80.5	484	6	ABK93497	Abk93497 Human bre
c 386	34	82.9	41498	11	ACN43858	Acn43858 Human gen	c 459	33	80.5	501	9	ACH39953	Ach39953 Human foe
c 387	34	82.9	42061	13	ABD33566	Abd33566 Human can	460	33	80.5	504	3	AAAC52575	Aac52575 Arabidops
c 388	34	82.9	56153	4	AA846793	Aa846793 Tumour su	c 461	33	80.5	506	4	ABA60510	AbA60510 Human foe
c 389	34	82.9	57036	14	ADZ113891	Adz113891 Human can	c 462	33	80.5	506	4	AAI40397	Aai40397 Probe #90
c 390	34	82.9	61020	4	AA846788	Aa846788 Tumour su	c 463	33	80.5	506	4	AAK34677	Aak34677 Human bon
c 391	34	82.9	64976	2	AAV21209_16	Continuation (17 o	c 464	33	80.5	506	4	ABS34449	AbS34449 Human liv
c 392	34	82.9	65608	6	ABL62910	Abi62910 Breast ca	465	33	80.5	515	13	ADX65718	Adx65718 Plant ful
c 393	34	82.9	65608	6	ABL64414	Abi64414 Stomach c	c 466	33	80.5	518	10	ADF50188	Adf50188 Salmonell
c 394	34	82.9	65608	6	ABL67668	Abi67668 Oesophagu	467	33	80.5	534	13	ACN61733	Acn61733 Cotton gy
c 395	34	82.9	67810	11	ACN45006	Acn45006 Human gen	c 468	33	80.5	543	14	ADM06405	Adm06405 Human gen
c 396	34	82.9	70251	9	ADA02606	Ada02606 Human IL2	c 469	33	80.5	551	13	ACN58298	Acn58298 Cotton gy
c 397	34	82.9	70251	10	ADB72344	Adb72344 Human IL2	c 470	33	80.5	555	4	AA551870	Aa551870 Staphyloc
c 398	34	82.9	70251	10	ADE95854	Ade95854 Human IL2	c 471	33	80.5	555	8	ABT15034	Abt15034 Pathogen
c 399	34	82.9	70271	14	ADZ12540	Adz12540 Human can	c 472	33	80.5	555	8	ACF73654	Acf73654 Staphyloc
c 400	34	82.9	90141	12	ADQ97867	Adq97867 Mouse can	c 473	33	80.5	558	4	AA855240	Aa855240 Staphyloc
c 401	34	82.9	90650	13	ADV35004	Adv35004 Murine CD	c 474	33	80.5	558	4	AA855424	Aa855424 Staphyloc
c 402	34	82.9	96592	9	ADA02678	Ada02678 Human TOP	c 475	33	80.5	558	4	AA854527	Aa854527 Staphyloc
c 403	34	82.9	96592	10	ADB72416	Adb72416 Human TIO	c 476	33	80.5	558	8	ACA20136	AcA20136 Prokaryoc
c 404	34	82.9	96592	10	ADE95926	Ade95926 Human TOP	477	33	80.5	561	13	ADQ52827	AdQ52827 Novel can
c 405	34	82.9	100696	14	ABE332384	AbE332384 Human gen	c 478	33	80.5	566	6	ABQ37052	Abq37052 Oligonucle
c 406	34	82.9	106373	13	ABD33737	Abd33737 Human can	c 479	33	80.5	566	6	ABQ37053	Abq37053 Oligonucle
c 407	34	82.9	108845	13	ABD32542	Abd32542 Mouse can	c 480	33	80.5	572	12	ACH78549	Ach78549 Human gen
c 408	34	82.9	110000	6	ABQ69245_09	Continuation (10 o	c 481	33	80.5	572	13	ACN60328	Acn60328 Cotton gy
c 409	34	82.9	110000	6	ABQ67197_08	Continuation (9 of	482	33	80.5	594	13	ACN60328	Acn60328 Cotton gy
c 410	34	82.9	110000	10	AB856454_07	Continuation (8 of	483	33	80.5	601	3	AA45692	Aa45692 Arabidops
c 411	34	82.9	110000	10	ACF67367_35	Continuation (36 o	c 484	33	80.5	601	6	ABN63038	Abn63038 Human can
c 412	34	82.9	110000	10	ACF67367_36	Continuation (37 o	c 485	33	80.5	603	3	AA78504	Aa78504 Plant SDF
c 413	34	82.9	110000	12	ADQ97050_2	Continuation (3 of	486	33	80.5	621	6	ABN70980	Abn70980 Streptoco
c 414	34	82.9	110000	13	ABD32806_2	Continuation (3 of	c 487	33	80.5	629	5	ABV55298	Abv55298 Human pro
c 415	34	82.9	110000	14	ADZ45062_07	Continuation (2 of	488	33	80.5	631	6	ABQ25623	Abq25623 Oligonucle
c 416	34	82.9	110000	14	ABE33174_01	Continuation (19 o	489	33	80.5	632	6	ABQ25622	Abq25622 Oligonucle
c 417	34	82.9	110000	14	ABE33175_18	Continuation (19 o	490	33	80.5	636	6	ABN67449	Abn67449 Streptoco
c 418	34	82.9	110000	14	ABE34201_18	Continuation (5 of	491	33	80.5	636	11	ACN86659	Acn86659 Breast ca
c 419	34	82.9	110000	14	ABE42737_04	Continuation (6 of	492	33	80.5	639	13	ADV83804	Adv83804 Streptoco
c 420	34	82.9	110000	14	ABE42737_05	Continuation (3 of	493	33	80.5	653	6	ABQ58517	Abq58517 Human col
c 421	34	82.9	118063	12	ADQ97180	Adq97180 Mouse can	494	33	80.5	670	14	ADY98173	Ady98173 Beta gluc
c 422	34	82.9	118063	12	ADQ97849	Adq97849 Mouse can	495	33	80.5	670	14	ADY98173	Ady98173 Beta gluc
c 423	34	82.9	118864	12	ADQ97818	Adq97818 Mouse can	496	33	80.5	675	6	ABZ32598	Abz32598 Candida a
c 424	34	82.9	119198	6	ABK83565	Abk83565 Human cDN	c 497	33	80.5	675	6	ABZ32598	Abz32598 Candida a
c 425	34	82.9	122557	12	ADH76849	Adh76849 Melanin-c	c 498	33	80.5	678	6	ABZ13655	Abz13655 Arabidops
c 426	34	82.9	127567	14	AEA61137	Aea61137 Human BCA	c 499	33	80.5	678	6	ABZ13655	Abz13655 Arabidops
c 427	34	82.9	137908	11	ADP65634	Adp65634 Human seq	c 500	33	80.5	678	8	ADA68620	Ada68620 Arabidops
c 428	34	82.9	155350	13	ADP33514	Adp33514 Murine ca	c 501	33	80.5	678	12	ADN72720	Adn72720 Thale cre
c 429	34	82.9	162450	3	AAZ86967	Aaz86967 Retinobla	c 502	33	80.5	681	5	ABV19321	Abv19321 Human pro
c 430	34	82.9	169865	12	ADQ97056	Adq97056 Human can	c 503	33	80.5	726	10	ADC90805	Adc90805 E. faeciu
c 431	34	82.9	174600	12	ADQ97520	Adq97520 Mouse can	c 504	33	80.5	754	6	ABQ13679	Abq13679 Oligonucle
c 432	34	82.9	215974	12	ADQ97523	Adq97523 Human can	c 505	33	80.5	754	6	ABQ13678	Abq13678 Oligonucle
c 433	34	82.9	334462	10	ADC24763	Adc24763 Human wil	c 506	33	80.5	765	10	ADD29986	Add29986 Plant yle
c 434	34	82.9	344548	11	ACN44070	Acn44070 Human gen	c 507	33	80.5	765	12	ADQ43680	AdQ43680 Plant tra
c 435	33	80.5	121	12	ADK93071	Adk93071 Polynucle	c 508	33	80.5	765	12	ADQ63006	AdQ63006 Transcrip
c 436	33	80.5	148	14	ABE01423	Aeb01423 MicroRNA	c 509	33	80.5	770	4	AAH07520	Aah07520 Human cDN
c 437	33	80.5	163	12	ADQ04656	Adq04656 Maize hom	c 510	33	80.5	780	4	AAH04155	Aah04155 Human cDN
c 438	33	80.5	249	12	ADP94581	Adp94581 Cotton ex	c 511	33	80.5	795	8	ACA50051	AcA50051 Prokaryot
c 439	33	80.5	259	3	AAA11367	Aaa11367 Human sec	c 512	33	80.5	811	13	ADX14572	Adx14572 Plant ful
c 440	33	80.5	282	6	ABN76336	Abn76336 Human ORF	c 513	33	80.5	841	6	ABQ17174	Abq17174 Oligonucle
c 441	33	80.5	282	13	ACN52277	Acn52277 Cotton an	c 514	33	80.5	841	6	ABQ17175	Abq17175 Oligonucle
c 442	33	80.5	284	13	ACN51293	Acn51293 Cotton an	515	33	80.5	873	5	AA869551	Aa869551 DNA encod

c 662	33	80.5	6125	6	ABL33612	Human imm	Ab133612	Human imm	c 735	33	80.5	26166	8	ABZ71491	Secreted
c 663	33	80.5	6125	6	ABK28277	DNA trans	Abk28277	DNA trans	c 736	33	80.5	26166	9	ADB91847	Human sec
c 664	33	80.5	6192	6	ABL32598	Human imm	Ab132598	Human imm	c 737	33	80.5	26166	10	ADC74629	Human sec
c 665	33	80.5	6192	6	ABL92210	Chemical	Ab192210	Chemical	c 738	33	80.5	26166	10	ADA57652	BAC fragm
c 666	33	80.5	6192	6	ABL54349	Chemical	Ab154349	Chemical	c 739	33	80.5	29023	10	AD885921	Human bra
c 667	33	80.5	6192	6	AD22311	Chemical	Ad22311	Chemical	c 740	33	80.5	29023	10	AD885922	Human bra
c 668	33	80.5	6418	6	ABL32323	Human imm	Ab132323	Human imm	c 741	33	80.5	29993	10	ADB37663	Human che
c 669	33	80.5	6418	6	AA561074	Human gen	Aa561074	Human gen	c 742	33	80.5	29993	10	ADB37661	Human che
c 670	33	80.5	6802	6	AA546281	Tumour su	Aa546281	Tumour su	c 743	33	80.5	31814	10	AA47150	Human ras
c 671	33	80.5	6802	6	ABK31178	Signal tr	Abk31178	Signal tr	c 744	33	80.5	32064	4	AA529832	Human cyt
c 672	33	80.5	6802	6	ABL70141	Chemical	Ab170141	Chemical	c 745	33	80.5	32221	4	AAK90119	Human dig
c 673	33	80.5	6802	6	AA561067	Human gen	Aa561067	Human gen	c 746	33	80.5	32221	5	AA539766	Genomic s
c 674	33	80.5	6816	4	AA546688	Tumour su	Aa546688	Tumour su	c 747	33	80.5	32221	5	ADB32726	Human nov
c 675	33	80.5	7143	6	ABL32982	Human imm	Ab132982	Human imm	c 748	33	80.5	34634	6	AAK31198	Human WKL
c 676	33	80.5	7304	6	ABL92268	Chemical	Ab192268	Chemical	c 749	33	80.5	34680	4	AAK81232	Human imm
c 677	33	80.5	7304	6	ABL49343	Human pol	Ab149343	Human pol	c 750	33	80.5	34680	4	AAK85743	Human imm
c 678	33	80.5	7747	4	AA507401	Arabidops	Ab507401	Arabidops	c 751	33	80.5	38342	4	AA546746	Tumour su
c 679	33	80.5	7892	6	ABK40056	Human che	Abk40056	Human che	c 752	33	80.5	38342	4	AA546745	Tumour su
c 680	33	80.5	8353	4	ABL15774	Drosophill	Ab115774	Drosophill	c 753	33	80.5	38342	4	AA546745	Tumour su
c 681	33	80.5	8547	6	ABK31204	Signal tr	Abk31204	Signal tr	c 754	33	80.5	38342	6	ABK31506	Signal tr
c 682	33	80.5	8547	6	ABL70171	Chemical	Ab170171	Chemical	c 755	33	80.5	38342	6	ABK31506	Signal tr
c 683	33	80.5	8547	6	AA561120	Human gen	Aa561120	Human gen	c 756	33	80.5	38342	6	ABK31507	Signal tr
c 684	33	80.5	9786	6	ABQ67082	Human ang	Abq67082	Human ang	c 757	33	80.5	38771	4	AAK81036	Human imm
c 685	33	80.5	10837	4	ABL29814	Drosophill	Ab129814	Drosophill	c 758	33	80.5	39567	8	AAK74053	Human imm
c 686	33	80.5	10837	4	ABL29814	Drosophill	Ab129814	Drosophill	c 759	33	80.5	39567	8	ABZ74429	Secreted
c 687	33	80.5	10907	4	ABL30104	Drosophill	Ab130104	Drosophill	c 760	33	80.5	39567	10	ABZ67986	Human sec
c 688	33	80.5	11046	6	ABK31536	Signal tr	Abk31536	Signal tr	c 761	33	80.5	41322	9	AA162633	Human CD3
c 689	33	80.5	11138	4	AAK69445	Human imm	Aak69445	Human imm	c 762	33	80.5	41587	4	AA164984	Moricella
c 690	33	80.5	11443	2	AAV52182	Streptoco	Aav52182	Streptoco	c 763	33	80.5	42424	4	ABL19930	Drosophill
c 691	33	80.5	11555	6	ABL32616	Human imm	Ab132616	Human imm	c 764	33	80.5	46846	11	ACN44526	Human gen
c 692	33	80.5	11555	6	AD28380	Human che	Ad28380	Human che	c 765	33	80.5	60935	13	ACN37224	Human per
c 693	33	80.5	14487	4	ABL20344	Drosophill	Ab120344	Drosophill	c 766	33	80.5	68355	8	ACF62737	Cancer ba
c 694	33	80.5	14542	6	ABK31234	Signal tr	Abk31234	Signal tr	c 767	33	80.5	68355	8	ADB20852	MRP1 base
c 695	33	80.5	14542	6	ABL70191	Chemical	Ab170191	Chemical	c 768	33	80.5	68355	10	ADB87941	Human UGT
c 696	33	80.5	14542	6	AA561147	Human gen	Aa561147	Human gen	c 769	33	80.5	68355	10	ADB86924	Human MDR
c 697	33	80.5	14768	4	ABL10258	Drosophill	Ab110258	Drosophill	c 770	33	80.5	68355	10	ADB92115	Human MDR
c 698	33	80.5	15413	4	AAK84002	Human imm	Aak84002	Human imm	c 771	33	80.5	68592	13	ABD33059	Mouse can
c 699	33	80.5	15479	6	ABK39964	Human che	Abk39964	Human che	c 772	33	80.5	70469	11	ACN44014_3	Continuation (4 of
c 700	33	80.5	15861	6	ABL32524	Human imm	Ab132524	Human imm	c 773	33	80.5	70768	6	AA141152	Wooden le
c 701	33	80.5	15951	6	ABL33680	Human imm	Ab133680	Human imm	c 774	33	80.5	76804	6	AB578942	E. coli C
c 702	33	80.5	15951	6	ABL34580	Human met	Ab134580	Human met	c 775	33	80.5	76804	10	ADH80509	Escherich
c 703	33	80.5	15951	6	ABL70373	Chemical	Ab170373	Chemical	c 776	33	80.5	77287	9	AA58261	Murine tu
c 704	33	80.5	15951	7	AD599841	Bisulphit	Ad599841	Bisulphit	c 777	33	80.5	77932	11	ADL27149	Mouse gen
c 705	33	80.5	15959	14	ADM94172	Staphyloc	Adw94172	Staphyloc	c 778	33	80.5	78028	9	ADA03071	Mouse MCG
c 706	33	80.5	16826	2	AAV74357	Staphyloc	Aav74357	Staphyloc	c 779	33	80.5	78028	9	ADA66355	Mouse MCG
c 707	33	80.5	16842	4	AA546412	Tumour su	Aa546412	Tumour su	c 780	33	80.5	78028	10	ADB72809	Mouse MCG
c 708	33	80.5	16842	6	ABK31419	Signal tr	Abk31419	Signal tr	c 781	33	80.5	79467	10	ADB72455	Mouse Nfa
c 709	33	80.5	16842	6	ABL70384	Chemical	Ab170384	Chemical	c 782	33	80.5	79467	10	ADB72455	Mouse Nfa
c 710	33	80.5	16842	6	AA561336	Human gen	Aa561336	Human gen	c 783	33	80.5	79467	10	ADB95965	Mouse Nfa
c 711	33	80.5	16842	10	ADB54136	Pretrate	Abd54136	Pretrate	c 784	33	80.5	80276	14	ADZ12738	Murine ca
c 712	33	80.5	16842	10	ADB54264	Pretrate	Abd54264	Pretrate	c 785	33	80.5	82952	6	ABN85766	Arabidops
c 713	33	80.5	16842	13	ADB89560	Oligonuc	Ad89560	Oligonuc	c 786	33	80.5	92562	10	ADC85284	Human ltk
c 714	33	80.5	16842	13	ADB89286	Oligonuc	Ad89286	Oligonuc	c 787	33	80.5	92563	9	ADA02804	Human ITK
c 715	33	80.5	16876	4	AAK65625	Human imm	Aak65625	Human imm	c 788	33	80.5	92563	10	ADB72542	Human ITK
c 716	33	80.5	16876	4	AAK65624	Human imm	Aak65624	Human imm	c 789	33	80.5	92563	12	ADM74399	Human car
c 717	33	80.5	17213	6	ABL33483	Human imm	Ab133483	Human imm	c 790	33	80.5	95400	12	ADP08388	Human lam
c 718	33	80.5	17293	6	ABL31171	Signal tr	Abk31171	Signal tr	c 791	33	80.5	95596	13	ADV87741	Streptoco
c 719	33	80.5	17293	6	ABL70126	Chemical	Ab170126	Chemical	c 792	33	80.5	95596	13	ADV78994	Streptoco
c 720	33	80.5	17293	6	AA561058	Human gen	Aa561058	Human gen	c 793	33	80.5	95596	9	ADA02864	Human ITP
c 721	33	80.5	17534	6	ABK40025	Human che	Abk40025	Human che	c 794	33	80.5	95596	10	ADB72602	Human ITP
c 722	33	80.5	18154	6	ABL32255	Human imm	Ab132255	Human imm	c 795	33	80.5	95596	10	ADC85343	Mouse ltp
c 723	33	80.5	18997	6	ABL32571	Human imm	Ab132571	Human imm	c 796	33	80.5	95596	12	ADM74459	Human car
c 724	33	80.5	18997	6	ABK33949	Human DNA	Abk33949	Human DNA	c 797	33	80.5	96599	9	ADA02981	Mouse Map
c 725	33	80.5	18997	8	ADA20353	Prostate	Ada20353	Prostate	c 798	33	80.5	96599	10	ADB72719	Mouse Map
c 726	33	80.5	18997	8	AD84160	Human ren	Ad84160	Human ren	c 799	33	80.5	96599	10	ADC85461	Mouse Map
c 727	33	80.5	19795	4	AD84160	Human ren	Ad84160	Human ren	c 800	33	80.5	96599	12	ADM74576	Murine ca
c 728	33	80.5	20001	13	ACN37214	Human per	Acn37214	Human per	c 801	33	80.5	98642	11	ACN44584	Mouse gen
c 729	33	80.5	20433	4	AAK65626	Human imm	Aak65626	Human imm	c 802	33	80.5	99960	3	AAZ50905	Human TBC
c 730	33	80.5	20933	6	ABQ67124	Human ang	Abq67124	Human ang	c 803	33	80.5	100554	11	ACN44624	Mouse gen
c 731	33	80.5	25000	12	ADI34524	Human GUC	Adi34524	Human GUC	c 804	33	80.5	101270	12	ADQ17814	Human sof
c 732	33	80.5	25020	12	ADO40235	S. agalac	Ado40235	S. agalac	c 805	33	80.5	106378	11	ACN44930	Human gen
c 733	33	80.5	26166	8	ADA1524	Human sec	Ada1524	Human sec	c 806	33	80.5	106645	13	ADT05645	Haemophil
c 734	33	80.5	26166	8	ACC50869	Human sec	Acc50869	Human sec	c 807	33	80.5	106938	13	ABD33432	Human can

808	33	80.5	106938	13	ADR67034	Adx67034 Human can	c 881	33	80.5	195917	12	ADQ20606	Adq20606 Human sof
809	33	80.5	107310	14	ADZ13456	Adz13456 Human can	c 882	33	80.5	198849	14	ADZ13007	Adz13007 Human can
c 810	33	80.5	109586	11	ACN43994	Acn43994 Human gen	c 883	33	80.5	200622	14	ABX39167	Abx39167 L. pneumo
c 811	33	80.5	110000	2	AA742063_06	Continuation (7 of)	c 884	33	80.5	201431	14	ABX35715	Abx35715 L. pneumo
812	33	80.5	110000	2	AAV21209_08	Continuation (9 of)	c 885	33	80.5	208648	8	ACF62735	Acf62735 Cancer ba
813	33	80.5	110000	2	AAV21209_08	Continuation (2 of)	c 886	33	80.5	208648	8	ACF62740	Acf62740 Cancer ba
814	33	80.5	110000	6	ABN71527_11	Continuation (12 of)	c 887	33	80.5	208648	8	ADZ20850	Adz20850 MRP1 base
815	33	80.5	110000	6	ABQ74964_5	Continuation (6 of)	c 888	33	80.5	208648	8	ADZ20855	Adz20855 MRP1 base
c 816	33	80.5	110000	6	ABQ74964_6	Continuation (7 of)	c 889	33	80.5	208648	10	ADB87944	Adb87944 Human UGT
c 817	33	80.5	110000	6	ABL57909_0	Abi57909 Human tra	c 890	33	80.5	208648	10	ADB87939	Adb87939 Human UGT
818	33	80.5	110000	6	ABA92787_1	Continuation (2 of)	c 891	33	80.5	208648	10	ADB96922	Adb96922 Human MDR
819	33	80.5	110000	6	ABA90521_05	Continuation (6 of)	c 892	33	80.5	208648	10	ADB96927	Adb96927 Human MDR
820	33	80.5	110000	6	ABA30341_15	Continuation (16 of)	c 893	33	80.5	208648	10	ADB92113	Adb92113 Human MDR
c 821	33	80.5	110000	6	ABA03041_17	Continuation (18 of)	c 894	33	80.5	208648	10	ADB92118	Adb92118 Human MDR
822	33	80.5	110000	10	ADP77343_03	Continuation (4 of)	c 895	33	80.5	220224	11	ACN44702	Acn44702 Human gen
823	33	80.5	110000	10	ADP77343_03	Continuation (19 of)	c 896	33	80.5	221500	14	ADX80720	Adx80720 Human neu
c 824	33	80.5	110000	10	ADP77343_18	Continuation (15 of)	c 897	33	80.5	227246	13	ADB33272	Adb33272 Human can
c 825	33	80.5	110000	13	ABD32804_3	Continuation (4 of)	c 898	33	80.5	228835	12	ADQ97421	Adq97421 Human can
826	33	80.5	110000	13	ABD32909_1	Continuation (2 of)	c 899	33	80.5	234882	14	ADZ13715	Adz13715 Human can
827	33	80.5	110000	13	ABD32923_6	Continuation (7 of)	c 900	33	80.5	241748	14	ADZ13116	Adz13116 Murine ca
828	33	80.5	110000	13	ADV81204_12	Continuation (13 of)	c 901	33	80.5	246940	12	ADQ59422	Adq59422 Human can
c 829	33	80.5	110000	14	ADZ42274_2	Continuation (3 of)	c 902	33	80.5	256190	13	ADB33276	Adb33276 Human can
c 830	33	80.5	110000	14	ABE339174_03	Continuation (4 of)	c 903	33	80.5	271990	12	ADN61228	Adn61228 Fertility
c 831	33	80.5	110000	14	ABE339175_29	Continuation (30 of)	c 904	33	80.5	271990	12	ADN61228	Adn61228 Radiat nu
c 832	33	80.5	110000	14	ABE42401_16	Continuation (17 of)	c 905	33	80.5	290040	14	ADV16961	Adv16961 Human pro
833	33	80.5	110000	14	ABE42737_02	Continuation (3 of)	c 906	33	80.5	300000	14	ADU92049	Adu92049 Human PAM
834	33	80.5	110000	14	ABE42737_03	Continuation (4 of)	c 907	33	80.5	300000	10	ADB86352	Adb86352 Human PTP
c 835	33	80.5	110000	14	ABE42737_14	Continuation (15 of)	c 908	33	80.5	300001	12	ADQ14076	Adq14076 Human pro
836	33	80.5	110218	11	ACN44744	Acn44744 Mouse gen	c 909	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
837	33	80.5	113515	2	ABL34174	Abi34174 Human imm	c 910	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
838	33	80.5	116624	2	AAV52850	AAV52850 Human eya	c 911	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
839	33	80.5	116704	11	ACN44818	Acn44818 Human gen	c 912	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 840	33	80.5	117829	12	ADQ97319	Adq97319 Human can	c 913	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 841	33	80.5	120670	12	ADQ97319	Adq97319 MSI-H car	c 914	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
842	33	80.5	130877	13	ABD33104	Abd33104 Human can	c 915	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 843	33	80.5	136284	6	ABK833575	Abk833575 Human cDN	c 916	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 844	33	80.5	136284	13	ADK52798	Adk52798 Drug ther	c 917	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
845	33	80.5	150201	14	ADZ13203	Adz13203 Human can	c 918	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 846	33	80.5	151826	3	AAF22291	AAF22291 BAC conta	c 919	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
847	33	80.5	154645	6	ADZ28763	Adz28763 Human AKA	c 920	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
848	33	80.5	158245	6	ADZ28763	Adz28763 Human AKA	c 921	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 849	33	80.5	158811	12	ADQ11901	Adq11901 Human sof	c 922	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
850	33	80.5	160755	4	AH88704	Ah88704 Human DNA	c 923	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
851	33	80.5	161334	11	ACN44334	Acn44334 Human gen	c 924	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
852	33	80.5	161425	4	AH02340	Ah02340 Human AKA	c 925	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
853	33	80.5	162025	4	AH02339	Ah02339 Human AKA	c 926	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
854	33	80.5	162025	6	ADZ28758	Adz28758 Human AKA	c 927	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
855	33	80.5	162025	6	ADZ28759	Adz28759 Human AKA	c 928	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
856	33	80.5	162025	13	ADZ5958	Adz5958 Human A-k	c 929	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
857	33	80.5	162025	13	ADZ5959	Adz5959 Human A-k	c 930	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
858	33	80.5	162025	13	ADZ5963	Adz5963 Human A-k	c 931	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
859	33	80.5	162025	13	ADZ5962	Adz5962 Human A-k	c 932	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
860	33	80.5	163319	3	AAF22306	AAF22306 Arabidops	c 933	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
861	33	80.5	167739	9	ADZ58258	Adz58258 Murine tu	c 934	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
862	33	80.5	171158	12	ADQ97894	Adq97894 Human can	c 935	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
863	33	80.5	177556	14	ABE71426	AbE71426 Human gen	c 936	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
864	33	80.5	177563	9	ACD28257	ACD28257 Mouse sol	c 937	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
865	33	80.5	181684	11	ACN44374	Acn44374 Human gen	c 938	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
866	33	80.5	183610	8	ACF62736	Acf62736 Cancer ba	c 939	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 867	33	80.5	183610	8	ADZ20851	Adz20851 MRP1 base	c 940	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 868	33	80.5	183610	10	ADB87940	Adb87940 Human UGT	c 941	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 869	33	80.5	183610	10	ADB96923	Adb96923 Human MDR	c 942	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 870	33	80.5	183610	10	ADB92114	Adb92114 Human MDR	c 943	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
871	33	80.5	186591	8	ACF62750	Acf62750 Cancer ba	c 944	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
872	33	80.5	186591	8	ADZ20869	Adz20869 MRP1 base	c 945	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
873	33	80.5	186591	10	ADB87958	Adb87958 Human UGT	c 946	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
874	33	80.5	186591	10	ADB96941	Adb96941 Human MDR	c 947	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
875	33	80.5	186591	10	ADB92132	Adb92132 Human MDR	c 948	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
c 876	33	80.5	192639	10	ADL13676	Adl13676 Osteoarth	c 949	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
877	33	80.5	193303	12	ADP13122	Adp13122 Hypermeth	c 950	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
878	33	80.5	193303	12	ADP13115	Adp13115 Hypermeth	c 951	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
879	33	80.5	193303	12	ADP13268	Adp13268 Hypermeth	c 952	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus
880	33	80.5	193303	12	ADP137261	Adp137261 Hypermeth	c 953	33	80.5	349980	5	AH41223	Ah41223 Pyrococcus

c 954 32 78.0 350 5 ABV01492 Human pro
 c 955 32 78.0 351 5 ABV10661 Human pro
 c 956 32 78.0 355 5 AAC29470 Human sec
 c 957 32 78.0 361 5 ABV59728 Human pro
 c 958 32 78.0 364 5 AAI91607 Human pol
 c 959 32 78.0 383 5 ABV13226 Human pro
 c 960 32 78.0 385 5 AAH94312 Human foe
 c 961 32 78.0 387 5 ABA08397 Human sec
 c 962 32 78.0 391 5 AAC19256 Human sec
 c 963 32 78.0 399 5 ADL38713 Human ova
 c 964 32 78.0 399 5 ADI73583 Human ova
 c 965 32 78.0 400 5 AAD30416 PCR prime
 c 966 32 78.0 401 5 ABV31828 Human pro
 c 967 32 78.0 401 5 ABV40793 Human pro
 c 968 32 78.0 401 13 ACP66053 Human SIR
 c 969 32 78.0 405 3 AAC54999 Arabidops
 c 970 32 78.0 405 4 AAS38325 Novel hum
 c 971 32 78.0 405 5 ABV04057 Human pro
 c 972 32 78.0 410 5 AAF66643 Novel hum
 c 973 32 78.0 411 10 ADC92292 E. faeciu
 c 974 32 78.0 413 8 ABX53360 Bovine ES
 c 975 32 78.0 416 2 AAQ59547 Human bra
 c 976 32 78.0 421 5 ABV34345 Human pro
 c 977 32 78.0 428 6 ABL67137 Thyroid c
 c 978 32 78.0 428 6 ABN93925 Gene #423
 c 979 32 78.0 430 5 AAF64751 Novel hum
 c 980 32 78.0 436 5 ACH48457 Human leu
 c 981 32 78.0 444 5 ABV05916 Human pro
 c 982 32 78.0 445 6 ABZ15270 Arabidops
 c 983 32 78.0 446 5 ADL39787 Human ova
 c 984 32 78.0 465 4 AAI12127 Probe #20
 c 985 32 78.0 465 4 ABA53833 Human foe
 c 986 32 78.0 465 4 ABA59194 Human foe
 c 987 32 78.0 465 4 AAI38941 Probe #76
 c 988 32 78.0 465 4 AAI33471 Probe #21
 c 989 32 78.0 465 4 AAI38971 Probe #76
 c 990 32 78.0 465 4 ABA27954 Probe #64
 c 991 32 78.0 465 4 ABA23578 Probe #20
 c 992 32 78.0 465 4 AAK33173 Human bon
 c 993 32 78.0 465 4 AAK27545 Human bon
 c 994 32 78.0 465 4 AAK33144 Human bon
 c 995 32 78.0 465 4 AAK02093 Human bra
 c 996 32 78.0 465 4 AAK07373 Human bra
 c 997 32 78.0 465 4 AAK07396 Human bra
 c 998 32 78.0 465 4 ABS27113 Human liv
 c 999 32 78.0 465 4 ABS32886 Human liv
 1000 32 78.0 465 4 ABS32919 Human liv

ALIGNMENTS

RESULT 1
 ID ADK11641 standard; DNA; 246 BP.
 XX AC ADK11641;
 XX DT 06-MAY-2004 (first entry)
 XX DE Breast cancer differentially expressed gene product #47.
 XX KW ds; cytostatic; gene therapy; DKP2p5661133 activity inhibitor;
 XX KW breast cancer; differential expression.
 XX OS Homo sapiens.
 XX XN WO2003057926-A1.
 XX 17-JUL-2003.
 XX PF 08-JAN-2003; 2003WO-US000657.
 XX PR 08-JAN-2002; 2002US-0345637P.

XX (CHIR) CHIRON CORP.
 XX Hansen R;
 XX WPI; 2003-577534/54.
 DR Inhibiting a cancerous phenotype of a cell, useful for treating breast
 PT cancer comprises contacting a cancerous mammalian cell with an agent for
 PT inhibition of DKP2p5661133 activity.
 XX Claim 30; SEQ ID NO 47; 257pp; English.
 PS The invention relates to a method of inhibiting a cancerous phenotype of
 CC a cell comprises contacting a cancerous mammalian cell with an agent for
 CC inhibition of DKP2p5661133 activity. The methods are useful for treating
 CC cancer, e.g. breast cancer. This sequence represents a gene product which
 CC is differentially expressed in breast cancer cells. The sequence can be
 CC used in the method of the invention.
 XX SQ Sequence 246 BP; 77 A; 49 C; 59 G; 61 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 87.9 Length: 246
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-8 (1-9) x ADK11641 (1-246)
 QY 1 AlaIlePheLeuLeuValleuTyrieu 9
 Db 15 GCTATTTTCCTCCTCGTTTGATTG 41
 RESULT 2
 ADU11677
 ID ADU11677 standard; DNA; 475 BP.
 AC ADU11677;
 DT 27-JAN-2005 (first entry)
 DE Solid tumour prognosis gene seqid 2116.
 XX cytostatic; gene therapy; expression profile; solid tumour;
 KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
 XX OS Unidentified.
 XX WO2004097052-A2.
 XX PD 11-NOV-2004.
 XX PF 29-APR-2004; 2004WO-US013587.
 XX PR 29-APR-2003; 2003US-0466067P.
 XX PR 23-JAN-2004; 2004US-0538246P.
 XX PA (AMHP) WYETH.
 XX PA (STRA/) STRAHS A.
 XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
 PI Immerman F, Dornier AJ;
 XX WPI; 2004-804779/79.
 DR A method, useful for prognosing and treating solid tumor, comprises
 PT comparing an expression profile of a gene expressed in peripheral blood
 PT mononuclear cells to a reference expression profile of a gene.
 XX Disclosure; Page; 111pp; English.

XX The invention describes a method comprising comparing an expression
 CC profile of at least one gene in a peripheral blood sample of a patient to
 CC at least one reference expression profile of the at least one gene, where
 CC the patient has a solid tumour, and each of the gene is differentially
 CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
 CC of patients as compared to PBMCs of a second class of patients, where
 CC both the first and second classes of patients have the solid tumour, and
 CC each of the first and second classes is a subcluster formed by an
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of
 CC a population of patients who have the solid tumour, and where the
 CC majority of the first class of patients has a first clinical outcome, and
 CC the majority of the second class of patients has a second clinical
 CC outcome. Also described are: a system comprising (i) a memory or a
 CC storage medium including data that represent an expression profile of at
 CC least one gene in a peripheral blood sample of a patient who has a solid
 CC tumour, (ii) at least another storage medium including data that
 CC represent at least one reference expression profile of the gene, (iii) a
 CC program capable of comparing the expression profile to the reference
 CC expression profile, and (iv) a processor capable of executing the
 CC program, where expression levels of the gene in peripheral blood
 CC mononuclear cells of patients who have the solid tumour correlate with
 CC clinical outcomes of the patients; and a nucleic acid or protein array
 CC comprising concentrated probes for solid tumour prognosis genes, where
 CC each of the solid tumour prognosis genes is differentially expressed in
 CC PBMCs of a first class of patients as compared to PBMCs of a second class
 CC of patients, where both the first and second classes of patients have a
 CC solid tumour, and where the first class of patients has a first clinical
 CC outcome, and the second class of patients has a second clinical outcome.
 CC The method, system, and array are useful for prognosing and treating
 CC solid tumours. This sequence represents a solid tumour prognosis gene of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 174 Length: 475
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-8 (1-9) x ADU11677 (1-475)

Qy 1 AlaIlePheLeuValLeuTyrLeu 9
 Db 381 GCTATTTCCTCCTCGTTTGATTG 407

RESULT 3
 AAA27060
 ID AAA27060 standard; DNA; 901 BP.

XX AAA27060;

XX 22-AUG-2000 (first entry)

XX Canine 5T4 tumour-associated antigen gene.

XX Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
 KW db.

XX Canis sp.

XX Key Location/Qualifiers

PH 1..858

FT /*tag= a

FT /product= "5T4 antigen"

FT misc_feature 61..74

FT /*tag= b

FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 135..146
 FT /*tag= c
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 207..216
 FT /*tag= d
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 277..290
 FT /*tag= e
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 351..361
 FT /*tag= f
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 422..436
 FT /*tag= g
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 497..511
 FT /*tag= h
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 572..583
 FT /*tag= i
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 644..653
 FT /*tag= j
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 714..723
 FT /*tag= k
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT misc_feature 784..801
 FT /*tag= l
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 FT WO200029428-A2.
 XX 25-MAY-2000.
 XX 18-NOV-1999; 99WO-GB003859.
 XX 18-NOV-1998; 98GB-00025303.
 XX 27-JAN-1999; 99GB-00001739.
 XX 30-JUL-1999; 99GB-00017995.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Carroll MW, Myers KA;
 XX WPI; 2000-387735/33.
 XX P-PSDB; AAY94351.
 XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.

XX PS Disclosure; Page 78-79; 79pp; English.

XX CC The present sequence encodes the canine 5T4 tumour-associated antigen (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in carcinomas but has a highly restricted expression pattern in normal adult tissues. It appears to be strongly correlated to metastasis in colorectal cancer and gastric cancer. 5T4 antigen may therefore be useful in tumour diagnosis, targeting and immunotherapy. Mice in which tumours had been induced were inoculated with a virus expression vector containing the human or murine 5T4 gene sequence. The 5T4 antigen was shown to be effective at eliciting an immunotherapeutic anti-tumour response. Both the nucleic acid encoding the antigen and the antigen itself can be used to elicit an immune response, preferably CTL or an antibody response in a subject

XX SQ Sequence 901 BP; 178 A; 246 C; 212 G; 153 T; 0 U; 112 Other;

Alignment Scores:

Pred. No.:	339	Length:	901
Score:	41.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	3	Indels:	0
DB:		Gaps:	0

US-10-774-176-8 (1-9) x AAA27060 (1-901)

QY 1 AlailePheLeuValLeuTyrlieu 9
|||||
669 GCCATCTTCCTAGCTGCTTTTGATTG 695

Db

RESULT 4

ABT07721

ID ABT07721 standard; DNA; 927 BP.

AC ABT07721;

XX 14-NOV-2002 (first entry)

XX Breast cancer-associated gene sequence 29.

XX Gene; ds; breast cancer; breast cancer-associated gene sequence; drug development; pharmacogenetics; biosensor development.

XX Unidentified.

XX WO200259377-A2.

XX 01-AUG-2002.

XX 24-JAN-2002; 2002WO-US002242.

XX 24-JAN-2001; 2001US-0263965P.

XX 02-FEB-2001; 2001US-0265928P.

XX 09-APR-2001; 2001US-00829472.

XX 09-APR-2001; 2001US-0282698P.

XX 04-MAY-2001; 2001US-0288590P.

XX 29-MAY-2001; 2001US-0294443P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC, Afar D;

XX WPI; 2002-583738/62.

XX N-PSDB; ABJ05564.

XX Detecting a breast cancer-associated transcript in a patient's cell, useful for diagnosing breast cancer, comprises contacting a biological sample with a polynucleotide that selectively hybridizes with breast cancer nucleic acids.

XX Claim 9; Page 372; 414pp; English.

XX CC The invention comprises a method of detecting a breast cancer-associated transcript in a cell from a patient. The method of the invention involves contacting a biological sample from the patient with a nucleotide that hybridises to one of the 69 breast cancer-associated gene sequences shown in the specification. The method of the invention is useful in the diagnosis or prognosis of breast cancer, and for detecting genes that are up or down-regulated in breast cancer cells. Genes identified by the method of the invention can be used in diagnostic purposes and also as targets for screening for therapeutic compounds that modulate breast cancer (e.g. hormones or antibodies). Identification of genes that are over or under expressed in breast cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure and biosensor development. DNA sequences ABT07693 - ABT07761 represent the 69 breast cancer-associated gene sequences of the invention

XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	349	Length:	927
Score:	41.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	6	Indels:	0
DB:		Gaps:	0

US-10-774-176-8 (1-9) x ABT07721 (1-927)

QY 1 AlailePheLeuValLeuTyrlieu 9
|||||
760 GCTATTTTCTCTGCTTTTGATTG 786

Db

RESULT 5

ABX76333

ID ABX76333 standard; DNA; 927 BP.

XX AC ABX76333;

XX 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #197.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema; antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis; small cell lung cancer; benign lesion; precancerous lesion; bronchitis; chronic obstructive pulmonary disease; hypersensitivity pneumonitis; interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

XX 10-MAY-2001; 2001US-0290492P.

XX 09-NOV-2001; 2001US-0339245P.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0334370P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX P-PSDB; ABU56604.

XX Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 XX Claim 22; Page 336; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 349 Length: 927
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-8 (1-9) x ABX76333 (1-927)

QY 1 AlaiPheLeuValLeuTyrLeu 9
 |||||
 Db 760 GCTATTTCCTCTGTTGTTGTTG 786

RESULT 6

ADB80503
 ID ADB80503 standard; DNA; 927 BP.

AC ADB80503;

XX 04-DEC-2003 (first entry)

XX Ovarian cancer-associated transcript #34.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; ds; gene.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1..927
 FT /*tag= a

FN WO2002102235-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

XX 12-APR-2002; 2002US-0372246P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

DR P-PSDB; ADB80504.

XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

XX Claim 10; Page 297; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.

XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 349 Length: 927
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-8 (1-9) x ADB80503 (1-927)

QY 1 AlaiPheLeuValLeuTyrLeu 9
 |||||
 Db 760 GCTATTTCCTCTGTTGTTGTTG 786

RESULT 7

ADN38723

ID ADN38723 standard; cDNA; 927 BP.

XX ADN38723;

XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

XX WO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-032464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-039775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
 XX P-PSDB; ADN38724.
 DR WPI; 2003-468649/44.
 DR P-PSDB; ADN38724.
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 CC Claim 8; SEQ ID NO 41; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as peptidias, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 349 Length: 927
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

US-10-774-176-8 (1-9) x ADN38723 (1-927)

Qy 1 AlallePheLeuValLeuTyrLeu 9
 |||||
 Db 760 GCTATTTCTCTCTGTTTGATTG 786

RESULT 8
 AAD56198
 ID AAD56198 standard; DNA; 973 BP.
 XX
 AC AAD56198;
 XX
 XX 07-AUG-2003 (first entry)
 DT
 DE Human LRRCAPS related DNA #5.

XX Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX WO2003035831-A2.
 XX
 PD 01-MAY-2003.
 XX
 XX 21-OCT-2002; 2002WO-US033540.
 XX
 XX 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX
 XX (EXEL-) EXELIXIS INC.
 PA
 XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX WPI; 2003-421410/39.
 DR
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 PS Example 5; Page 74-75; 99pp; English.
 XX
 CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX
 SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 367 Length: 973
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-8 (1-9) x AAD56198 (1-973)

Qy 1 AlallePheLeuValLeuTyrLeu 9
 |||||
 Db 775 GCTATTTCTCTCTGTTTGATTG 801

RESULT 9
 ABV99349
 ID ABV99349 standard; DNA; 1156 BP.
 XX
 AC ABV99349;
 XX
 XX 27-JAN-2003 (first entry)
 DT
 XX Human NOV8a coding sequence.
 DE
 XX Human; anti-HIV; cytostatic; antidiabetic; antisthmatic; cachexia; AIDS;
 KW antinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
 KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;

KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

XX Homo sapiens.

OS WO200272771-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US007288.

XX 08-MAR-2001; 2001US-0274101P.

XX 08-MAR-2001; 2001US-0274194P.

XX 08-MAR-2001; 2001US-0274281P.

XX 08-MAR-2001; 2001US-0274322P.

XX 09-MAR-2001; 2001US-0274849P.

XX 12-MAR-2001; 2001US-0275235P.

XX 13-MAR-2001; 2001US-0275578P.

XX 13-MAR-2001; 2001US-0275579P.

XX 13-MAR-2001; 2001US-0275601P.

XX 14-MAR-2001; 2001US-0276000P.

XX 16-MAR-2001; 2001US-0276776P.

XX 19-MAR-2001; 2001US-0276994P.

XX 20-MAR-2001; 2001US-0277239P.

XX 20-MAR-2001; 2001US-0277321P.

XX 20-MAR-2001; 2001US-0277327P.

XX 20-MAR-2001; 2001US-0277338P.

XX 21-MAR-2001; 2001US-0277791P.

XX 22-MAR-2001; 2001US-0277833P.

XX 23-MAR-2001; 2001US-0278152P.

XX 26-MAR-2001; 2001US-0278894P.

XX 27-MAR-2001; 2001US-0278999P.

XX 27-MAR-2001; 2001US-0279036P.

XX 28-MAR-2001; 2001US-0279344P.

XX 30-MAR-2001; 2001US-0279995P.

XX 30-MAR-2001; 2001US-0280233P.

XX 02-APR-2001; 2001US-0280802P.

XX 02-APR-2001; 2001US-0280822P.

XX 02-APR-2001; 2001US-0280900P.

XX 04-APR-2001; 2001US-0281194P.

XX 13-APR-2001; 2001US-0283675P.

XX 30-APR-2001; 2001US-0287424P.

XX 02-MAY-2001; 2001US-0288066P.

XX 03-MAY-2001; 2001US-0288342P.

XX 15-MAY-2001; 2001US-0291190P.

PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX (CURA-) CURAGEN CORP.
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VI, Pochart PF;
PI Zhong M;
XX WPI; 2002-732824/79.
DR P-PSDB; ABP70071.
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
PT disorders, and asthma.
XX Claim 16; Page 114-115; 619pp; English.
XX The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, haematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 439 Length: 1156
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-8 (1-9) x ABV99349 (1-1156)
QY 1 AlarilePheLeuValLeuTyrLeu 9
DB 991 GCTATTTCTCTCTGTTTATTG 1017
RESULT 10
ABK87175
ID ABK87175 standard; cDNA; 1260 BP.
XX AC ABK87175;
XX 07-OCT-2002 (first entry)
DT cDNA encoding feline oncofoetal leucine-rich glycoprotein, 574.
DE Feline; cat; oncofoetal leucine-rich glycoprotein; 574; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX

OS Felis sp.
XX Key Location/Qualifiers
XX CDS 1..1260
FT /*tag= a
FT /product= "5T4 protein"
XX WO200238612-A2.
XX 16-MAY-2002.
XX 13-NOV-2001; 2001WO-GB005004.
XX 13-NOV-2000; 2000WO-GB004317.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX Myers K, Drury N, Carroll M;
XX WPI; 2002-557449/59.
XX P-PSDB; AAU98694.
XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
XX polypeptide, useful in preparation of vaccine for treating and/or
XX preventing cancer in a subject, preferably a dog or cat.
XX Claim 4; Page 68; 60pp; English.
XX The present invention relates to the isolation of canine and feline
XX oncofetal leucine-rich glycoproteins known as 5T4, and the
XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
XX a significant proportion of tumours. The sequences of the invention are
XX useful in a pharmaceutical composition for the prevention and/or
XX treatment of tumours or other diseases associated with cell
XX proliferation, infections, and inflammatory conditions in animals,
XX preferably dogs or cats. The compositions may also be used for cancer
XX immunotherapy in these animals. The sequences of the invention may also
XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
XX measurement and localisation of 5T4 in extracts of plasma, urine,
XX tissues, and in cell culture media. Antibodies specific for the 5T4
XX protein are useful for isolating foetal cells from maternal blood. The
XX isolation process may form part of a diagnostic method e.g. the foetal
XX cells may then be subject to biochemical or genetic sampling used for
XX testing foetal abnormalities, or to determine the sex of the foetus(es).
XX The present sequence encodes feline 5T4 protein
SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 480 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x ABK87175 (1-1260)
QY 1 AlallePheLeuValLeuTyrlieu 9
DB 1099 GCCATTTCCTACTGGTTTGACTTG 1125

RESULT 11
ADB97513
ID ADB97513 standard; DNA; 1260 BP.
XX ADB97513;
XX 04-DEC-2003 (first entry)
XX Feline 5T4 antigen DNA.

XX Major Histocompatibility Complex class I peptide epitope; MHC;

KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
KW cytostatic; cancer; feline; gene; ds.
OS Unidentified.
XX Key Location/Qualifiers
XX CDS 1..1260
FT /*tag= a
FT /product= "Feline 5T4 antigen protein"
XX WO2003068816-A1.
XX 21-AUG-2003.
XX 13-FEB-2003; 2003WO-GB000670.
XX 13-FEB-2002; 2002GB-00003419.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX Carroll M, Kingsman S, Redchenko I;
XX WPI; 2003-637141/60.
XX P-PSDB; ADB97520.
XX New major histocompatibility complex class I peptide epitopes from human
XX 5T4 tumor-associated antigen, useful for preventing and/or treating a
XX disease, particularly cancer.
XX Disclosure; Page 67; 73pp; English.
XX The invention relates to a novel Major Histocompatibility Complex (MHC)
XX class I peptide epitope of the 5T4 antigen. The invention further
XX provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
XX sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
XX epitope; a vector system capable of delivering the 5T4 epitope nucleic
XX acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
XX 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
XX comprising the above; a method for treating and/or preventing a disease
XX in a subject by administering the vaccine; an agent capable of binding
XX specifically to the 5T4 epitope and/its encoding nucleic acid; a method
XX comprising detecting the presence of the 5T4 epitope or its encoding
XX nucleic acid in a subject; and a T cell line or clone capable of
XX specifically recognising the 5T4 epitope in conjunction with an MHC class
XX I molecule. The 5T4 epitope has cytostatic activity. The vaccine
XX comprising the 5T4 epitope or its encoding nucleic acid and the vector
XX system or cell is useful in the prevention and/or treatment of a disease,
XX particularly cancer. The detection method is useful for diagnosing or
XX monitoring the progression of a cancerous disease, and for detecting the
XX presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
XX is useful in the manufacture of a medicament for treating and/or
XX preventing a disease. This polynucleotide sequence represents the feline
XX 5T4 antigen coding DNA of the invention.
XX Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 480 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-8 (1-9) x ADB97513 (1-1260)
QY 1 AlallePheLeuValLeuTyrlieu 9
DB 1099 GCCATTTCCTACTGGTTTGACTTG 1125

RESULT 12
ADB97452
ID ADB97452 standard; DNA; 1260 BP.

```

XX AC ADB97452;
XX DT 04-DEC-2003 (first entry)
XX DE DNA encoding feline 5T4 protein.
XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
XX KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
XX OS Unidentified.
XX FH Key
XX FT CDS 1..1260
XX FT /*tag= a
XX FT /product= "Feline 5T4 antigen protein"
XX PN WO2003068815-A2.
XX PD 21-AUG-2003.
XX PF 13-FEB-2003; 2003WO-GB000618.
XX PR 13-FEB-2002; 2002GB-00003420.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll M, Harrop R, Kingsman S;
XX DR WPI; 2003-663795/62.
XX DR P-PSDB; ADB97455.
XX PT New Major Histocompatibility Complex class II peptide epitope of 5T4,
XX PT useful for manufacturing a medicament for diagnosing, preventing and/or
XX PT treating a disease, e.g. cancer.
XX PI Disclosure; Page 49; 63pp; English.
XX CC The invention relates to a Major Histocompatibility Complex (MHC) class
XX CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
XX CC clone has a cytostatic activity, as it is useful in manufacturing a
XX CC medicament for preventing and/or treating a disease, particularly cancer.
XX CC The methods are useful for detecting T-cells capable of specifically
XX CC recognising a peptide epitope in conjunction with an MHC molecule, for
XX CC diagnosing or monitoring the progression of a cancerous disease, or for
XX CC detecting the presence of a peptide or nucleic acid using an agent. The
XX CC MHC class II peptide epitope of the invention can be used in gene therapy
XX CC or as part of a vaccine. This polynucleotide sequence represents the DNA
XX CC coding for the feline 5T4 protein.
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 480 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-8 (1-9) x ADB97452 (1-1260)
QY 1 AlallePheLeuValLeuTyrLeu 9
Db 1099 GCCATTTCCTTACTGGTTTGACTTG 1125
RESULT 13
AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX AC AAA27058;
XX DT 22-AUG-2000 (first entry)

```

```

XX DE Human 5T4 tumour-associated antigen gene.
XX KW Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX KW ds.
XX OS Homo sapiens.
XX PN WO200029428-A2.
XX PD 25-MAY-2000.
XX PF 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MW, Myers KA;
XX DR WPI; 2000-387735/33.
XX PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX PT response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the human 5T4 tumour-associated antigen
XX CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The 5T4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject
XX SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 481 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-8 (1-9) x AAA27058 (1-1263)
QY 1 AlallePheLeuValLeuTyrLeu 9
Db 1102 GCTATTTCCTCTCGTTTGATTG 1128
RESULT 14
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX AC AAF89736;
XX DT 23-JUL-2001 (first entry)
XX DE Nucleotide sequence of canine 5T4 protein.
XX KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
XX KW hypersensitivity; autoimmune disease; central nervous system disorder;
XX KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
XX KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
XX KW Helicobacter-related disease; immune disorder; ss.

```

XX OS Canis sp.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..1263
 XX FT /*tag= a
 XX FT /product= "5T4"
 XX PN WO200136486-A2.
 XX PD 25-MAY-2001.
 XX PF 13-NOV-2000; 2000WO-GB004317.
 XX PR 18-NOV-1999; 99WO-GB003859.
 XX PR 15-PEB-2000; 2000GB-00003527.
 XX PR 02-MAR-2000; 2000GB-00005071.
 XX XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PA Kingman A, Kingman SM, Bebbington CR, Carroll MW, Ellard FM;
 XX PI Myers KA;
 XX PI WPI; 2001-343805/36.
 XX DR P-PSDB; AAB83839.
 XX XX Use of single chain antibody capable of recognizing a disease associated
 XX PT molecule for manufacturing a medicament for preventing and/or treating a
 XX FT disease condition associated with disease associated molecule.
 XX XX Disclosure; Fig 26; 118pp; English.
 XX CC The specification describes the use of a single chain antibody (ScFv),
 XX CC which is capable of recognizing a disease associated molecule in the
 XX CC manufacture of a medicament for the prevention and treatment of a disease
 XX CC condition. The ScFv antibody is useful in the manufacture of a
 XX CC medicament, for affecting a disease in vivo, for preparing a
 XX CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
 XX CC treatment of a disease. The ScFv antibody is also useful for treating
 XX CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
 XX CC diseases, cancers, central nervous system disorders including Parkinson's
 XX CC diseases, pericardial diseases, cardiopulmonary diseases, cardiovascular
 XX CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
 XX CC related diseases, and other immune disorders. The present sequence
 XX CC encodes a 5T4 protein, which is used to produce ScFv of the invention
 XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 481 Length: 1263
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-8 (1-9) x AAF89736 (1-1263)
 QY 1 AlallePheLeuValLeuTyrLeu 9
 DB 1102 GCCATCTTCCTACTGGTTTGTATTG 1128

RESULT 15
 ABK87174
 ID ABK87174 standard; cDNA; 1263 BP.
 AC ABK87174;
 XX 07-OCT-2002 (first entry)
 XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
 XX Canine, dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;

KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX OS Canis sp.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..1263
 XX FT /*tag= a
 XX FT /product= "5T4 protein"
 XX PN WO200238612-A2.
 XX PD 16-MAY-2002.
 XX PF 13-NOV-2001; 2001WO-GB005004.
 XX PR 13-NOV-2000; 2000WO-GB004317.
 XX XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PA Myers K, Drury N, Carroll M;
 XX PI WPI; 2002-557449/59.
 XX DR P-PSDB; AAU98693.
 XX XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 XX PT polypeptide, useful in preparation of vaccine for treating and/or
 XX FT preventing cancer in a subject, preferably a dog or cat.
 XX XX Claim 1; Page 67; 68pp; English.
 XX CC The present invention relates to the isolation of canine and feline
 XX CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 XX CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 XX CC a significant proportion of tumours. The sequences of the invention are
 XX CC useful in a pharmaceutical composition for the prevention and/or
 XX CC treatment of tumours or other diseases associated with cell
 XX CC proliferation, infections, and inflammatory conditions in animals,
 XX CC preferably dogs or cats. The compositions may also be used for cancer
 XX CC immunotherapy in these animals. The sequences of the invention may also
 XX CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 XX CC measurement and localisation of 5T4 in extracts of plasma, urine,
 XX CC tissues, and in cell culture media. Antibodies specific for the 5T4
 XX CC protein are useful for isolating foetal cells from maternal blood. The
 XX CC isolation process may form part of a diagnostic method e.g. the foetal
 XX CC cells may then be subject to biochemical or genetic sampling used for
 XX CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 XX CC The present sequence encodes canine 5T4 protein
 XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 481 Length: 1263
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x ABK87174 (1-1263)
 QY 1 AlallePheLeuValLeuTyrLeu 9
 DB 1102 GCCATCTTCCTACTGGTTTGTATTG 1128

RESULT 16
 AAA27059
 ID AAA27059 standard; DNA; 1281 BP.
 XX AAA27059;
 XX 22-AUG-2000 (first entry)

```

XX DE Mouse 5T4 tumour-associated antigen gene.
XX KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX OS ds.
XX XX Mus musculus.
XX FN WO200029428-A2.
XX PD 25-MAY-2000.
XX XX 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MM, Myers KA;
XX DR WPI; 2000-387735/33.
XX PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX PS response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the mouse 5T4 tumour-associated antigen
XX CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The 5T4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject. The
XX CC present sequence appears in GenBank at accession number AJ012160
XX SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 488 Length: 1281
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-8 (1-9) x AAA27059 (1-1281)
Qy 1 Alal1ePheLeuValLeuTyrLeu 9
Db 1120 GCTATTTTCCTCCTCGTTTGTATTG 1146

RESULT 17
AAD56199
ID AAD56199 standard; DNA; 1331 BP.
XX AC
XX AAD56199;
XX AC
XX 07-AUG-2003 (first entry)
XX DT
XX Human LRRCAPS related DNA #6.
XX DE
XX Human; p53 pathway; Leucine rich repeat capricious related protein;
XX KW LRRCAPS; cancer; gene therapy; ds.
XX OS Homo sapiens.

```

```

XX WO2003035831-A2.
XX PD 01-MAY-2003.
XX PF 21-OCT-2002; 2002WO-US033540.
XX PR 22-OCT-2001; 2001US-0338733P.
XX PR 15-FEB-2002; 2002US-0357600P.
XX PR 01-MAR-2002; 2002US-0361196P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
XX PI Francis-Lang H, Friedman L;
XX DR WPI; 2003-421410/39.
XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
XX PT comprises contacting an assay system comprising a purified leucine rich
XX PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX PS Disclosure; Page 75-76; 99pp; English.
XX CC The invention relates to a method of identifying a candidate p53 pathway
XX CC modulating agent. The method involves contacting an assay system
XX CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX CC polypeptide or nucleic acid or its fragment with a test agent and
XX CC detecting a test agent-biased activity, where a difference between the
XX CC test agent-biased activity and the reference activity identifies the test
XX CC agent as a candidate p53 pathway modulating agent. The method is useful
XX CC for identifying a candidate p53 pathway-modulating agent for preparing a
XX CC composition for diagnosing or treating cancer. The invention is useful in
XX CC gene therapy. The present sequence is human LRRCAPS related DNA
XX SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 508 Length: 1331
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-8 (1-9) x AAD56199 (1-1331)
Qy 1 Alal1ePheLeuValLeuTyrLeu 9
Db 1132 GCTATTTTCCTCCTCGTTTGTATTG 1158

RESULT 18
ADJ56299
ID ADJ56299 standard; cDNA; 2020 BP.
XX AC
XX ADJ56299;
XX DT
XX 06-MAY-2004 (first entry)
XX DE
XX Human cDNA differentially expressed in MYCN activated cells SeqID 105.
XX KW human; differential expression; transactivator; proto-oncogene;
XX KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
XX KW MYCN activated cell.
XX OS Homo sapiens.
XX PN US2003119009-A1.
XX XX
XX PD 26-JUN-2003.
XX XX
XX 25-FEB-2002; 2002US-00084817.
XX XX

```


PR 23-FEB-2001; 2001US-0270784P.
 XX (STUA/) STUART S G.
 PA (NUCH/) NUCHTERN J G.
 PA (PLON/) PLON S E.
 PA (SHOH/) SHOHET J M.
 XX
 XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
 PI WPI, 2003-635698/60.
 XX
 XX New genes regulated by MYCN activation, useful in gene therapy,
 PT particularly for treating a subject with e.g. neuroblastoma or other
 PT cancers, or for diagnosing, staging or monitoring the treatment of the
 PT cancer.
 XX
 XX Claim 1; SEQ ID NO 105; 27pp; English.
 XX
 XX This invention relates to novel isolated cDNAs that are differentially
 CC expressed in MYCN activated cells. Specifically, it refers to
 CC polynucleotide sequences that exhibit differential expression patterns in
 CC cells activated by the transactivator MYCN, where MYCN is a proto-
 CC oncogene that is amplified in neuroblastoma cells and is common in small
 CC cell lung cancers. The present invention describes these cDNA molecules
 CC as useful for in hybridisation assays to detect expression of nucleic
 CC acids (or complementary nucleic acids) in a present in a given sample, as
 CC well as for screening assays by identifying molecules or compounds that
 CC specifically bind the cDNA as a ligand and modulate function or activity.
 CC Accordingly, these compositions exhibit cytostatic activity and can also
 CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
 CC that is differentially expressed in MYCN activated cells, given in an
 CC exemplification of the invention. NOTE: This sequence does not appear in
 CC the printed specification but has been obtained in electronic format from
 CC the US Patent Office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.
 XX
 XX Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 783 Length: 2020
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-8 (1-9) x ADJ56299 (1-2020)

Qy 1 AlaiIlePheLeuValIeuTyIleu 9
 |||||
 Db 1172 GCTATTTTCCTCGTGGTTTGATTG 1198

RESULT 19
 ACC51052
 ID ACC51052 standard; cDNA; 2053 BP.

XX ACC51052;

XX 12-JUN-2003 (first entry)

XX Human bladder cancer associated cDNA sequence SEQ ID NO:192.

XX Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

XX WO2003003906-A2.

XX 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US021338.

XX 03-JUL-2001; 2001US-0302814P.

PR 03-AUG-2001; 2001US-0310099P.
 PR 08-NOV-2001; 2001US-0343705P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Mack DH, Aziz N;
 XX
 XX WPI, 2003-201532/19.
 DR P-PSDB; ABR48236.

XX Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT bladder cancer-associated polynucleotide or antibody.

XX Claim 6; Page 296; 307pp; English.

XX The present invention describes a method for detecting a bladder cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridises to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications

XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 797 Length: 2053
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-8 (1-9) x ACC51052 (1-2053)

Qy 1 AlaiIlePheLeuValIeuTyIleu 9
 |||||
 Db 1186 GCTATTTTCCTCGTGGTTTGATTG 1212

RESULT 20
 ABX76332
 ID ABX76332 standard; DNA; 2053 BP.

XX ABX76332;

XX 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #196.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

PR 01-MAR-2002; 2002US-0361196P.
 XX (EXEL-) EXELIXIS INC.
 XX Belvin M, Schleithoff L, Plozman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX WPI; 2003-421410/39.
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 XX Disclosure; Page 76-77; 99pp; English.
 XX
 CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 797 Length: 2053
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-8 (1-9) x AAD56200 (1-2053)
 Qy 1 AlallePheLeuValLeuTyrLeu 9
 Db 1186 GCTATTTTCCTCGTTTGTATTG 1212
 RESULT 23
 ADN38721
 ID ADN38721 standard; cDNA; 2053 BP.
 XX
 AC ADN38721;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.
 XX
 DE Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 XX 13-NOV-2002; 2002WO-US036810.
 XX
 PF 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Heverzi PA;
 PI Mack DR, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX WPI; 2003-468649/44.
 DR P-PSDB; ADN38722.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 XX Claim 8; SEQ ID NO 39; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides, and
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 797 Length: 2053
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0
 US-10-774-176-8 (1-9) x ADN38721 (1-2053)
 Qy 1 AlallePheLeuValLeuTyrLeu 9
 Db 1186 GCTATTTTCCTCGTTTGTATTG 1212
 RESULT 24
 ADL06473
 ID ADL06473 standard; cDNA; 2053 BP.
 XX
 XX ADL06473;
 AC
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #53.
 XX
 KW Human; tumour-associated antigenic target; TAT; cell death; tumour;

KW cancer; cytostatic; gene; ss.
XX Homo sapiens.
OS WO2004016225-A2.
XX 26-FEB-2004.
XX PD
XX 19-AUG-2003; 2003WO-US025892.
XX PF
XX 19-AUG-2003; 2002US-0404809P.
PR 21-AUG-2002; 2002US-0405645P.
PR 23-SEP-2002; 2002US-0413192P.
PR 15-OCT-2002; 2002US-0413008P.
PR 15-NOV-2002; 2002US-0426847P.
PR 02-JUL-2003; 2003US-0484959P.
XX PA (GETH) GENENTECH INC.
XX PI Desauvage FJ, Prantz G, Hillan KJ, Polakis P, Polson A, Smith V;
PI Spencer SD, Wu TD, Zhang Z;
XX WPI; 2004-257144/24.
DR P-PSDB; ADL06552.
XX New antibody that binds to a tumor-associated antigenic target (TAT)
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating cancer.
XX Claim 1; SEQ ID NO 53; 319pp; English.
XX CC The present invention relates to the isolation of human tumour-associated
XX antigenic target (TAT) polynucleotide and polypeptide sequences. Also
XX disclosed is an antibody that binds to a TAT polypeptide. The antibody is
XX a monoclonal antibody, an antibody fragment, a chimeric antibody or a
XX humanised antibody. It is conjugated to a growth inhibitory agent. It is
XX produced in bacteria or in CHO cells and induces death of a cell to which
XX it binds. The antibody is useful for preparing a composition for
XX diagnosing or treating tumours and cancer. The present sequence
XX represents a human TAT cDNA sequence of the invention.
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 797 Length: 2053
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-8 (1-9) x ADL06473 (1-2053)
QY 1 AlallePheLeuValLeuTyrLeu 9
Db 1186 GCTATTTCCTCCTCGTTTGATTG 1212

RESULT 25
ADN03961
ID ADN03961 standard; cDNA; 2053 BP.
XX AC
XX ADN03961;
DT 01-JUL-2004 (first entry)
XX DE Antipsoriatic cDNA sequence #180.
XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX OS Homo sapiens.
XX WO2004028479-A2.
XX

PD 08-APR-2004.
XX 25-SEP-2003; 2003WO-US030907.
XX 25-SEP-2002; 2002US-0414006P.
XX (GETH) GENENTECH INC.
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX WPI; 2004-305105/28.
DR P-PSDB; ADN03962.
XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX Claim 1; SEQ ID NO 355; 3069pp; English.
XX CC The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.
XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 797 Length: 2053
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-8 (1-9) x ADN03961 (1-2053)
QY 1 AlallePheLeuValLeuTyrLeu 9
Db 1186 GCTATTTCCTCCTCGTTTGATTG 1212

RESULT 26
ADR25444
ID ADR25444 standard; DNA; 2053 BP.
XX AC
XX ADR25444;
DT 21-OCT-2004 (first entry)
XX DE Breast cancer prognosis marker #1305.
XX ds; breast cancer; prognosis; gene expression; diagnosis.
XX OS Homo sapiens.
XX WO2004065545-A2.
XX PD 05-AUG-2004.
XX PF 15-JAN-2004; 2004WO-US001100.
XX PR 15-JAN-2003; 2003US-00342887.
XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX PA (NECA-) NETHERLANDS CANCER INST.
XX Van't Veer LJ, He Y;
XX WPI; 2004-593473/57.
XX Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of

PT five genes in a cell sample taken from patient, to control levels.
XX Disclosure; SEQ ID NO 1305; 226pp; English.
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 797 Length: 2053
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-8 (1-9) x ADR25444 (1-2053)
Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 1186 GCTATTTTCCTCGTGGTTTGATTTG 1212

RESULT 27
ACN38510
ID ACN38510 standard; cDNA; 2053 BP.
XX
AC ACN38510;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA103471, SEQ ID NO:2070.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
WO2004030615-A2.
XX
PN 15-APR-2004.
XX
PD 29-SEP-2003; 2003WO-US028547.
XX
PF 02-OCT-2002; 2002US-0414971P.
XX
PR (GETH) GENENTECH INC.
XX
PA Wu TD, Zhang Z, Zhou Y;
XX
FI WPI, 2004-347921/32.
XX
DR P-PDB; ABM80804.
XX
DR New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 2070; 7273pp; English.
XX

CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 797 Length: 2053
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-8 (1-9) x ACN38510 (1-2053)
Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 1186 GCTATTTTCCTCGTGGTTTGATTTG 1212

RESULT 28
ADV35098
ID ADV35098 standard; cDNA; 2053 BP.
XX
AC ADV35098;
XX
DT 10-FEB-2005 (first entry)
XX
DE Human cDNA of an exemplary efficacy gene for BAD SeqID174.
XX
KW human; ss; multi-parameter high throughput screening; MPHTS;
KW disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
KW bipolar affective disorder; BAD; autism; Parkinson's;
KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.
XX
OS Homo sapiens.
XX
PN US2003096264-A1.
XX
PD 22-MAY-2003.
XX
PF 18-JUN-2002; 2002US-00175523.
XX
PR 18-JUN-2001; 2001US-0299151P.
PR 07-SEP-2001; 2001US-0317828P.
PR 25-SEP-2001; 2001US-0325150P.
PR 14-NOV-2001; 2001US-033047P.
PR 18-JAN-2002; 2002US-0349936P.
PR 04-MAR-2002; 2002US-0361834P.
XX
XX (PSYC-) PSYCHIATRIC GENOMICS INC.
PA
PI Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
PI Palfreyman M, Rajan P;
XX
XX WPI; 2004-118903/12.
DR

XX Identifying a compound that can treat disease or disorders, such as, a
 PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
 PT determining the expression of one or more efficacy genes in a cell
 PT contacted with the test compound.
 XX
 PS Example 6; SEQ ID NO 174; 39pp; English.
 PS
 XX This invention relates to a novel screening method identified as a multi-
 CC parameter high throughput screening (MPHTS) assay. Specifically, it
 CC refers to an assay that utilises the disease signature of a plurality of
 CC specific genes associated with a particular disease, and identifies
 CC differential expression between those cells taken from individuals
 CC affected by that disease and those that are not affected. The present
 CC invention then describes the screening of candidate pharmaceutical
 CC compounds to identify those that have a potential therapeutic benefit for
 CC the treatment of neuropsychiatric and neurodegenerative disorders
 CC including schizophrenia, bipolar affective disorder (BAD) and autism, as
 CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
 CC of this invention exhibit various activities including neuroleptic,
 CC nootropic, antianimic and antidepressant. Furthermore, the screening
 CC method used in MPHTS will be automated, such that a large number of test
 CC compounds may be rapidly screened with a minimal amount of labour and
 CC effort. This polynucleotide is a human cDNA sequence of a gene that is
 CC differentially expressed in the presence of a therapeutic compound and
 CC represents an exemplary efficacy gene for bipolar affective disorder,
 CC given in an exemplification of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 797 Length: 2053
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-8 (1-9) x ADV35098 (1-2053)

Qy 1 AlalleleLeuValLeuTyrlLeu 9
 Db 1186 GCTATTTCCTCTGTTTGTATTG 1212

RESULT 29

AA87175
 ID AAS87175 standard; cDNA; 2338 BP.

XX AAS87175;

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22979.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

PA Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG22988.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX

PS Claim 1; SEQ ID NO 22979; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 912 Length: 2338
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0

US-10-774-176-8 (1-9) x AAS87175 (1-2338)

Qy 1 AlalleleLeuValLeuTyrlLeu 9

Db 1443 GCTATTTCCTCTGTTTGTATTG 1469

RESULT 30

AAK94253

ID AAK94253 standard; cDNA; 2359 BP.

XX AAK94253;

XX 06-NOV-2001 (first entry)

XX Human full-length cDNA, SEQ ID NO: 2864.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-00114089.

XX 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 XX P-PSDB; AAM93334.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX
 XX Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX
 SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 920 Length: 2359
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-8 (1-9) x AAK94253 (1-2359)
 QY 1 AlalilePheLeuValIeuTyrLeu 9
 DB 1525 GCTATTTCCTCCTGGTTTGTATTG 1551
 RESULT 31
 ID ADL30831 standard; cDNA; 2359 BP.
 XX
 AC ADL30831;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Full length human cDNA clone SeqID 2864.
 XX
 KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.
 XX
 OS Homo sapiens.
 XX
 XX EPI396543-A2.
 XX
 XX 10-MAR-2004.
 XX
 XX 07-JUL-2000; 2003EP-00025638.
 XX
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 FA
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 XX WPI; 2004-204755/20.
 DR P-PSDB; ADL30832.

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX
 XX Example 1; SEQ ID NO 2864; 1340pp; English.
 XX
 CC This invention relates to a novel primers useful for synthesising full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.
 XX
 SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 920 Length: 2359
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-8 (1-9) x ADL30831 (1-2359)
 QY 1 AlalilePheLeuValIeuTyrLeu 9
 DB 1525 GCTATTTCCTCCTGGTTTGTATTG 1551
 RESULT 32
 AAK94254
 ID AAK94254 standard; cDNA; 2361 BP.
 XX
 AC AAK94254;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human full-length cDNA, SEQ ID NO: 2866.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX
 OS Homo sapiens.
 XX
 XX EPI130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-00114089.
 XX
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 XX WPI; 2001-524255/58.
 DR P-PSDB; AAM93334.
 XX
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX
 XX Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX

CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO

XX
 SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 921 Length: 2361
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-774-176-8 (1-9) x AAK94254 (1-2361)

QY 1 AlAliePheLeuValLeuTyrLeu 9
 Db 1527 GCTATTTCCTCTGCTTTTGATTG 1553

RESULT 33

AD126162

ID AD126162 standard; cDNA; 2361 BP.

XX AC AD126162;

DT 22-APR-2004 (first entry)

DE Human cDNA encoding protein that promotes STAT6 activation #64.

XX ss; gene; human; signal transducer and activator of transcription 6;

KW STAT6; immunogen; STAT6 activation; allergy; inflammation;

KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;

KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;

KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

XX Homo sapiens.

OS WO2003104277-A2.

PN 18-DEC-2003.

PD 05-JUN-2003; 2003WO-JP007123.

PP 06-JUN-2002; 2002JP-00164257.

PR 26-DEC-2002; 2002US-0385912P.

PR 27-DEC-2002; 2002JP-00377326.

PR 15-MAY-2003; 2002US-0436467P.

PR 16-MAY-2003; 2003JP-00137505.

XX (ASAH) ASAMI KASEI KK.

PA Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX WPI; 2004-122214/12.

XX P-PSDB; AD126163.

DR New signal transducer and activator of transcription 6 activation

XX PT promoting purified protein, for diagnosing and treating disease

PT associated with activation/inhibition of transcription factor e.g.

XX diabetes and cancer.

XX Claim 4; SEQ ID NO 127; 1368pp; English.

XX The invention relates to a purified protein promoting signal transducer

CC and activator of transcription 6 activation (STAT6). The protein is

CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.

XX SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 921 Length: 2361
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-8 (1-9) x AD126162 (1-2361)

QY 1 AlAliePheLeuValLeuTyrLeu 9
 Db 1527 GCTATTTCCTCTGCTTTTGATTG 1553

RESULT 34

AD130833

ID ADL30833 standard; cDNA; 2361 BP.

XX AC ADL30833;

XX 20-MAY-2004 (first entry)

XX Full length human cDNA clone SeqID 2866.

XX human; medicine; signal transduction; glycoprotein; transcription;

KW oligo-capping method; ss; gene.

XX Homo sapiens.

XX EP1396543-A2.

XX 10-MAR-2004.

XX 07-JUL-2000; 2003EP-00025638.

XX 08-JUL-1999; 99JP-00194486.

XX 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183865.

XX 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

XX P-PSDB; ADL30834.

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full

XX length human cDNAs.

XX Example 1; SEQ ID NO 2866; 1340pp; English.

XX This invention relates to a novel primers useful for synthesising full

XX length cDNA molecules that encode human proteins. Specifically, it refers

XX to secretory or membrane proteins that are potential therapeutic agents/

XX target molecules in the field of medicine, and in particular genes

XX encoding proteins that are associated with signal transduction,

XX glycoproteins and transcription. The present invention describes a method

XX for efficiently cloning a full length human cDNA from both the 5' and 3'

XX ends using the oligo-capping method. This polynucleotide sequence is a

XX full length human cDNA clone of the invention.

SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 921 Length: 2361

Score: 41.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 12 Gaps: 0

US-10-774-176-8 (1-9) x ADL30833 (1-2361)

Qy 1 AlallePheLeuValLeuTyrieu 9

Db 1527 GCTATTTTCCTCCTCGTTTGTATTG 1553

RESULT 35

AD126160

ID AD126160 standard; cDNA; 2557 BP.

XX AC AD126160;

XX 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #63.

ss; gene; human; signal transducer and activator of transcription 6;

STAT6; immunogen; STAT6 activation; allergy; inflammation;

autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;

Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;

systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

OS Homo sapiens.

XX WO2003104277-A2.

PN 18-DEC-2003.

XX 05-JUN-2003; 2003WO-JP007123.

XX 05-JUN-2002; 2002JP-00164257.

PR 06-JUN-2002; 2002US-0385912P.

PR 26-DEC-2002; 2002JP-00377326.

PR 27-DEC-2002; 2002US-0436467P.

PR 15-MAY-2003; 2003JP-00137505.

PR 16-MAY-2003; 2003US-0470836P.

XX (ASAH) ASAMI KASEI KK.

XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX WPI; 2004-122214/12.

XX P-PSDB; AD126161.

XX New signal transducer and activator of transcription 6 activation

XX promoting purified protein, for diagnosing and treating disease

XX associated with activation/inhibition of transcription factor e.g.

XX diabetes and cancer.

XX

Claim 4; SEQ ID NO 125; 1368pp; English.

PS The invention relates to a purified protein promoting signal transducer

XX and activator of transcription 6 activation (STAT6). The protein is

XX useful for the producing an antibody, which involves administering the

XX protein or its epitope-bearing fragments to a non-human animal as an

XX antigen. The nucleic acid is useful for diagnosing a disease or

XX susceptibility to a disease related to expression or activity of the

XX protein. A transformant expressing the protein is useful for screening

XX compounds which inhibit or promote STAT6 activation. A transformant

XX expressing the protein is useful for producing a pharmaceutical

XX composition. Compositions, antibodies and antisense molecules are useful

XX for the treating a disease associated with STAT6 activation such as

XX allergic diseases, inflammation, autoimmune diseases, diabetes,

XX hyperlipidaemia, infections disease and cancers. Compositions are useful

XX for treating disease associated with STAT6 activation and/or prevention

XX of Th1 hyperactive diseases. Compositions are also useful in rheumatoid

XX arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,

XX allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,

XX viral hepatitis and AIDS. The protein has efficient promoting STAT6

XX activity. The protein or nucleic acid is effectively useful for screening

XX compounds for treating and preventing disease associated with excessive

XX activation or inhibition of STAT6. The present sequence represents a

XX human cDNA encoding a protein which promotes STAT6 activation.

SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1e+03 Length: 2557

Score: 41.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 12 Gaps: 0

US-10-774-176-8 (1-9) x AD126160 (1-2557)

Qy 1 AlallePheLeuValLeuTyrieu 9

Db 1675 GCTATTTTCCTCCTCGTTTGTATTG 1701

RESULT 36

AD126158

ID AD126158 standard; cDNA; 2557 BP.

XX AC AD126158;

XX 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #62.

ss; gene; human; signal transducer and activator of transcription 6;

STAT6; immunogen; STAT6 activation; allergy; inflammation;

autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;

Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;

systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

OS Homo sapiens.

XX WO2003104277-A2.

PN 18-DEC-2003.

XX 05-JUN-2003; 2003WO-JP007123.

XX 05-JUN-2002; 2002JP-00164257.

PR 06-JUN-2002; 2002US-0385912P.

PR 26-DEC-2002; 2002JP-00377326.

PR 27-DEC-2002; 2002US-0436467P.

PR 15-MAY-2003; 2003JP-00137505.

PR 16-MAY-2003; 2003US-0470836P.

XX

PA (ASAH) ASahi KASEI KK.

XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX WPI; 2004-122214/12.

DR P-PSDB; ADI26159.

XX New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
XX diabetes and cancer.

PS Claim 4; SEQ ID NO 123; 1368pp; English.

XX The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.

XX Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1e+03 Length: 2557
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-8 (1-9) x ADI26158 (1-2557)

Qy 1 AlallePheLeuValLeuTyrLeu 9

Db 1675 GCTATTTCCTCCTCGTTTGTATTG 1701

RESULT 37

AA20580/c

ID AAX20580 standard; DNA; 19142 BP.

XX AAX20580;

XX 05-MAY-1999 (first entry)

XX Polynucleotide sequence from the genome of *Treponema pallidum*.

XX *Treponema pallidum* infection; syphilis; *Borrelia* infection; animal;

XX enzyme production; db.

XX *Treponema pallidum*.

XX WO9859034-A2.

XX 30-DEC-1998.

XX 23-JUN-1998; 98WO-US013041.

XX 24-JUN-1997; 97US-0050667P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Fraser CM;

XX WPI; 1999-081273/07.

XX New isolated *Treponema pallidum* nucleic acids - used to develop products
PT for the detection, diagnosis, characterisation, prevention and therapy of
PT *T. pallidum* infections, particularly syphilis.

XX Claim 1; Page 608-619; 1150pp; English.

XX AAX20500-21243 represent polynucleotide sequences from the genome of
CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,
CC characterisation, prevention and therapy for *T. pallidum* infections,
CC particularly syphilis. They can also be used for detecting diseases
CC related to *Borrelia* infections in animals, and for the production of
CC biosynthetic products such as enzymes

XX Sequence 19142 BP; 4629 A; 5539 C; 4716 G; 4238 T; 0 U; 20 Other;

Alignment Scores:
Pred. No.: 1.22e+04 Length: 19142
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 97.6% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-8 (1-9) x AAX20580 (1-19142)

Qy 1 AlallePheLeuValLeuTyrLeu 9

Db 14194 GCGGTATTTTTCGTTTATATCTC 14168

RESULT 38

ADH84583

ID ADH84583 standard; DNA; 204 BP.

XX ADH84583;

XX 22-APR-2004 (first entry)

XX *Enterococcus faecalis* polynucleotide #2468.

XX *Enterococcus faecalis* infection; transcription regulatory element;
XX antibacterial; gene; ds.

XX *Enterococcus faecalis*.

XX US6617156-B1.

XX 09-SEP-2003.

XX 13-AUG-1998; 98US-00134000.

XX 15-AUG-1997; 97US-0055778P.

XX (DOUC/) DOUCETTE-STAMM L A.

XX (BUSH/) BUSH D.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2003-895394/82.

XX P-PSDB; ADH87988.

XX New nucleic acid comprising a sequence encoding an *Enterococcus faecalis*
PT polypeptide, useful for preparing a composition for diagnosing or
PT treating *E. faecalis* infection.

XX

PS Disclosure; SEQ ID NO 2468; 193pp; English.

XX The invention relates to Enterococcus faecalis polynucleotides and polypeptides. The invention also relates to a recombinant expression vector comprising a polynucleotide operably linked to a transcription regulatory element, a cell comprising a recombinant vector, a method for producing an E. faecalis polypeptide, an isolated nucleic acid comprising a sequence not given in the specification, a recombinant vector comprising the nucleic acid and a cell comprising the recombinant vector. The polynucleotides can be used to detect the presence of E. faecalis in a sample. The sequences are useful for preparing a composition for diagnosing or treating Enterococcus faecalis infection. This sequence represents an E. faecalis polynucleotide of the invention.

XX Sequence 204 BP; 68 A; 24 C; 38 G; 74 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 375 Length: 204
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-8 (1-9) x ADH84583 (1-204)

Qy 1 AlalpheLeuValleuTyrlieu 9

Db 22 OCCATTTTCTCTCGTTCCTTTTGTG 48

RESULT 39

ABL06215/C
ID ABL06215 standard; cDNA; 1058 BP.

XX ABL06215;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 13127.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX MPI; 2001-656860/75.

XX P-PSDB; ABB62112.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

XX Claim 1; SEQ ID NO 13127; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1058 BP; 298 A; 288 C; 259 G; 213 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,098+03 Length: 1058
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-8 (1-9) x ABL06215 (1-1058)

Qy 2 IlePheLeuValleuTyrlieu 9

Db 260 ATTTCCTCTGCTGCTATCTT 237

RESULT 40

AAC46422

ID AAC46422 standard; DNA; 1421 BP.

XX AAC46422;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 50080.

XX Hybridisation assay; Genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 06-MAY-1999; 99US-0132487P.

XX 07-MAY-1999; 99US-0132863P.

XX 11-MAY-1999; 99US-0134256P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 18-MAY-1999; 99US-0134370P.

XX 18-MAY-1999; 99US-0134768P.

XX 19-MAY-1999; 99US-0134941P.

XX 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999;	99US-0135353P.
PR 24-MAY-1999;	99US-0135629P.
PR 25-MAY-1999;	99US-0136021P.
PR 27-MAY-1999;	99US-0136392P.
PR 28-MAY-1999;	99US-0136782P.
PR 01-JUN-1999;	99US-0137222P.
PR 03-JUN-1999;	99US-0137528P.
PR 04-JUN-1999;	99US-0137502P.
PR 07-JUN-1999;	99US-0137724P.
PR 08-JUN-1999;	99US-0138094P.
PR 10-JUN-1999;	99US-0138540P.
PR 10-JUN-1999;	99US-0138847P.
PR 14-JUN-1999;	99US-0139119P.
PR 16-JUN-1999;	99US-0139452P.
PR 17-JUN-1999;	99US-0139453P.
PR 17-JUN-1999;	99US-0139492P.
PR 18-JUN-1999;	99US-0139454P.
PR 18-JUN-1999;	99US-0139847P.
PR 18-JUN-1999;	99US-0139455P.
PR 18-JUN-1999;	99US-0139456P.
PR 18-JUN-1999;	99US-0139457P.
PR 18-JUN-1999;	99US-0139458P.
PR 18-JUN-1999;	99US-0139459P.
PR 18-JUN-1999;	99US-0139460P.
PR 18-JUN-1999;	99US-0139461P.
PR 18-JUN-1999;	99US-0139462P.
PR 18-JUN-1999;	99US-0139463P.
PR 18-JUN-1999;	99US-0139750P.
PR 18-JUN-1999;	99US-0139763P.
PR 21-JUN-1999;	99US-0139817P.
PR 22-JUN-1999;	99US-0139899P.
PR 23-JUN-1999;	99US-0140353P.
PR 23-JUN-1999;	99US-0140354P.
PR 24-JUN-1999;	99US-0140693P.
PR 28-JUN-1999;	99US-0140823P.
PR 29-JUN-1999;	99US-0140991P.
PR 30-JUN-1999;	99US-0141287P.
PR 01-JUL-1999;	99US-0141842P.
PR 01-JUL-1999;	99US-0142154P.
PR 02-JUL-1999;	99US-0142055P.
PR 06-JUL-1999;	99US-0142390P.
PR 08-JUL-1999;	99US-0142803P.
PR 09-JUL-1999;	99US-0142920P.
PR 12-JUL-1999;	99US-0142977P.
PR 13-JUL-1999;	99US-0143542P.
PR 19-JUL-1999;	99US-0144331P.
PR 19-JUL-1999;	99US-0144332P.
PR 19-JUL-1999;	99US-0144333P.
PR 19-JUL-1999;	99US-0144334P.
PR 19-JUL-1999;	99US-0144335P.
PR 20-JUL-1999;	99US-0144352P.
PR 20-JUL-1999;	99US-0144632P.
PR 20-JUL-1999;	99US-0144884P.
PR 21-JUL-1999;	99US-0144814P.
PR 21-JUL-1999;	99US-0145086P.
PR 21-JUL-1999;	99US-0145088P.
PR 22-JUL-1999;	99US-0145085P.
PR 22-JUL-1999;	99US-0145087P.
PR 22-JUL-1999;	99US-0145089P.
PR 22-JUL-1999;	99US-0145192P.
PR 23-JUL-1999;	99US-0145145P.
PR 23-JUL-1999;	99US-0145218P.
PR 23-JUL-1999;	99US-0145224P.
PR 26-JUL-1999;	99US-0145276P.
PR 27-JUL-1999;	99US-0145913P.
PR 27-JUL-1999;	99US-0145918P.
PR 27-JUL-1999;	99US-0145919P.
PR 28-JUL-1999;	99US-0145951P.
PR 02-AUG-1999;	99US-0146386P.

PR 02-AUG-1999;	99US-0146388P.
PR 02-AUG-1999;	99US-0146389P.
PR 03-AUG-1999;	99US-0147038P.
PR 04-AUG-1999;	99US-0147204P.
PR 04-AUG-1999;	99US-0147302P.
PR 05-AUG-1999;	99US-0147192P.
PR 05-AUG-1999;	99US-0147260P.
PR 06-AUG-1999;	99US-0147303P.
PR 06-AUG-1999;	99US-0147416P.
PR 09-AUG-1999;	99US-0147493P.
PR 09-AUG-1999;	99US-0147935P.
PR 10-AUG-1999;	99US-0148171P.
PR 11-AUG-1999;	99US-0148319P.
PR 12-AUG-1999;	99US-0148341P.
PR 13-AUG-1999;	99US-0148565P.
PR 13-AUG-1999;	99US-0148684P.
PR 16-AUG-1999;	99US-0149368P.
PR 17-AUG-1999;	99US-0149175P.
PR 18-AUG-1999;	99US-0149426P.
PR 20-AUG-1999;	99US-0149722P.
PR 20-AUG-1999;	99US-0149723P.
PR 20-AUG-1999;	99US-0149929P.
PR 23-AUG-1999;	99US-0149902P.
PR 23-AUG-1999;	99US-0149930P.
PR 25-AUG-1999;	99US-0150566P.
PR 26-AUG-1999;	99US-0150884P.
PR 27-AUG-1999;	99US-0151065P.
PR 27-AUG-1999;	99US-0151066P.
PR 27-AUG-1999;	99US-0151080P.
PR 30-AUG-1999;	99US-0151303P.
PR 31-AUG-1999;	99US-0151438P.
PR 01-SEP-1999;	99US-0151930P.
PR 07-SEP-1999;	99US-0152363P.
PR 10-SEP-1999;	99US-0153070P.
PR 13-SEP-1999;	99US-0153758P.
PR 15-SEP-1999;	99US-0154018P.
PR 16-SEP-1999;	99US-0154039P.
PR 20-SEP-1999;	99US-0154779P.
PR 22-SEP-1999;	99US-0155139P.
PR 23-SEP-1999;	99US-0155486P.
PR 24-SEP-1999;	99US-0155659P.
PR 28-SEP-1999;	99US-0156458P.
PR 29-SEP-1999;	99US-0156596P.
PR 04-OCT-1999;	99US-0157117P.
PR 05-OCT-1999;	99US-0157753P.
PR 06-OCT-1999;	99US-0157865P.
PR 07-OCT-1999;	99US-0158029P.
PR 08-OCT-1999;	99US-0158232P.
PR 12-OCT-1999;	99US-0158369P.
PR 13-OCT-1999;	99US-0159293P.
PR 13-OCT-1999;	99US-0159294P.
PR 13-OCT-1999;	99US-0159295P.
PR 14-OCT-1999;	99US-0159329P.
PR 14-OCT-1999;	99US-0159330P.
PR 14-OCT-1999;	99US-0159331P.
PR 14-OCT-1999;	99US-0159637P.
PR 14-OCT-1999;	99US-0159638P.
PR 18-OCT-1999;	99US-0159584P.
PR 21-OCT-1999;	99US-0160741P.
PR 21-OCT-1999;	99US-0160767P.
PR 21-OCT-1999;	99US-0160768P.
PR 21-OCT-1999;	99US-0160770P.
PR 21-OCT-1999;	99US-0160814P.
PR 21-OCT-1999;	99US-0160815P.
PR 22-OCT-1999;	99US-0160980P.
PR 22-OCT-1999;	99US-0160981P.
PR 22-OCT-1999;	99US-0160989P.
PR 25-OCT-1999;	99US-0161404P.
PR 25-OCT-1999;	99US-0161405P.
PR 25-OCT-1999;	99US-0161406P.
PR 26-OCT-1999;	99US-0161359P.
PR 26-OCT-1999;	99US-0161360P.
PR 26-OCT-1999;	99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores: 2.82e+03 Length: 1421
Pred. No.: 37.00 Matches: 7
Score: 100.0% Conservative: 2
Percent Similarity: 77.8% Mismatches: 0
Best Local Similarity: 90.2% Indels: 0
Query Match: 3 Gaps: 0
DB:

US-10-774-176-8 (1-9) x AAC46422 (1-1421)

Qy 1 AlailePheLeuValLeuTyrlieu 9
Db 371 GCGATTTCCTGTAATCGTCTACTTA 397

RESULT 41

AAC36467

ID AAC36467 standard; DNA; 1422 BP.

XX AAC36467;

AC AAC36467 (first entry)

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13915.

XX Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway; metabolic pathway;
promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 23-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 14-MAY-1999; 99US-0134370P.

XX 18-MAY-1999; 99US-0134768P.

XX 19-MAY-1999; 99US-0134941P.

XX 20-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.

```
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0158294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160711P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2.82e+03 Length: 1422
Score: 37.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-8 (1-9) x AAC36467 (1-1422)
Qy 1 AlaIlePheIeuLeuValIeuTyrLeu 9
Db 371 GCGATTTTCTTGCTAATCGTCTACTTA 397

RESULT 42
ADT15316
ID ADT15316 standard; cDNA; 1689 BP.
XX
AC ADT15316;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant cDNA, Seq ID 642.
XX
KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW drought tolerance; disease resistance; galactomannan production;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW lignin production; extreme osmotic condition tolerance;
KW pathogen resistance; pest resistance; yield improvement; seed oil yield;
KW seed protein yield.
XX
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.
XX
PR 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
PI Kovalic DK;
XX
WPI; 2004-757369/74.
XX
PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
PS Claim 1; SEQ ID NO 642; 14pp; English.
XX
CC The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
```

CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.

XX
SQ Sequence 1689 BP; 430 A; 371 C; 326 G; 486 T; 0 U; 76 Other;

Alignment Scores:
Pred. No.: 3.37e+03 Length: 1689
Score: 37.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-8 (1-9) x ADT15316 (1-1689)

Qy 1 AlailePheLeuValLeuTyrLeu 9
|||
Db 371 GCGATTTCTGCTAATGCTACTTA 397

RESULT 43

ABL06214
ID ABL06214 standard; cDNA; 3298 BP.

XX ABL06214;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 13124.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUN-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR P-P8DB; ABB62111.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

PS Claim 1; SEQ ID NO 13124; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 3298 BP; 893 A; 751 C; 688 G; 966 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.76e+03 Length: 3298
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-8 (1-9) x ABL06214 (1-3298)

Qy 2 IlePheLeuValLeuTyrLeu 9
|||
Db 1799 ATTTCCTGCTGCTGTAUUTT 1822

RESULT 44

ABL33308

ID ABL33308 standard; DNA; 6224 BP.

XX ABL33308;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1281.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytosolic; neurotropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
ds.

XX Homo sapiens.

OS WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.

PS Claim 1; SEQ ID NO 1281; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences

CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX

SQ Sequence 6224 BP; 1471 A; 222 C; 1628 G; 2899 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.31e+04 Length: 6224
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x ABL33308 (1-6224)

QY 2 IlePheLeuValLeuTyrLeu 9
DB 4116 ATATTTTATTAGTTTATATTTA 4139

RESULT 45

ABL54355
ID ABL54355 standard; DNA; 6224 BP.

XX AC ABL54355;

XX DT 29-JUL-2002 (first entry)

XX DE Chemically treated apoptosis gene #28.

XX KW Apoptosis; HIV; Bloom syndrome; cardiopathy; neurodegenerative disorder;
XX KW Herpes simplex virus; renal ischaemia; amyotrophic lateral sclerosis;
XX KW cancer; ds.

XX OS Unidentified.

XX PN WO200177164-A2.

XX PD 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-EP003969.

XX PR 06-APR-2000; 2000DE-01019058.

XX PR 07-APR-2000; 2000DE-01019173.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2002-017444/02.

XX Chemically modified sequences of genes associated with apoptosis are
XX useful to determine methylation patterns of genomic DNA samples for
XX diagnosis of associated diseases such as cancer.

XX PS Claim 1; Seq ID #55; 24pp; English.

XX This invention relates to chemically pre-treated DNA of genes associated
XX with apoptosis. The nucleic acids are used to allocate patients for
XX specific therapy for HIV infection, Bloom syndrome, cardiopathy, aging,
XX neurodegenerative disorders, Herpes simplex virus infection, renal
XX ischaemia, amyotrophic lateral sclerosis, solid tumours and cancers. This
XX nucleotide sequence represents a chemically treated apoptosis gene. Even
XX SEQ ID numbers are the complementary DNA strands to the odd SEQ ID
XX numbers. The sequence data for this patent is not represented in the
XX printed specification but is based on information supplied by the
XX European patent office

SQ Sequence 6224 BP; 1471 A; 222 C; 1628 G; 2899 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.31e+04 Length: 6224
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x ABL54355 (1-6224)

QY 2 IlePheLeuValLeuTyrLeu 9
DB 4116 ATATTTTATTAGTTTATATTTA 4139

RESULT 46

AAAX13336

ID AAAX13336 standard; DNA; 32768 BP.

XX AC AAAX13336;

XX DT 19-MAR-1999 (first entry)

XX DE Enterococcus faecalis genome contig SEQ ID NO:399.

XX KW Enterococcus faecalis; contig; detection; Enterococcal infection;

XX KW vaccine; attenuation; computer readable medium; ds.

XX OS Enterococcus faecalis.

XX PN WO9850555-A2.

XX PD 12-NOV-1998.

XX PF 04-MAY-1998; 98WO-US008985.

XX PR 06-MAY-1997; 97US-0044031P.

XX PR 16-MAY-1997; 97US-0046655P.

XX PR 14-NOV-1997; 97US-0066009P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Kunsch CA, Dillon PJ, Barash SC;

XX DR WPI; 1999-045171/04.

XX PT New isolated Enterococcus faecalis polynucleotides and polypeptides -
XX used to develop products for the detection of Enterococcus and for use in
XX vaccines for prevention or attenuation of Enterococcus infection.

XX PS Claim 1; Page 1592-1609; 2084pp; English.

XX A computer readable medium has been developed which has recorded on it
XX 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
XX AAAX12938 to AAAX13919 represent these nucleotide sequences which are
XX primary nucleotide sequences, also known as contigs. The computer-based
XX system can identify fragments of the Enterococcus faecalis genome with
XX commercial importance. The products can be used to detect the presence of
XX Enterococcus faecalis in samples. They can also be used for diagnosing
XX Enterococcal infection in an animal and monitoring progression of
XX disease, and for identifying agents which can be used to modulate the
XX growth or pathogenicity of Enterococcus faecalis, or another related
XX organism, in vivo or in vitro. In particular the polypeptides encoded by
XX the Enterococcus faecalis nucleotide sequences can be used in vaccines to
XX prevent or attenuate an Enterococcal infection

SQ Sequence 32768 BP; 10797 A; 5960 C; 7083 G; 8921 T; 0 U; 7 Other;

Alignment Scores:

Pred. No.: 7.35e+04 Length: 32768
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1

Best Local Similarity: 88.9% Mismatches: 0
Query Match: 90.2% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-8 (1-9) x AAX13336 (1-32768)

Qy 1 AlallePheLeuValleuTyrIeu 9
Db 30913 GCCATTTTCTCTCGTCTCTTTTG 30939

RESULT 47
ABS99131
ID ABS99131 standard; DNA; 32768 BP.

AC ABS99131;
XX 18-DEC-2002 (first entry)
XX Enterococcus faecalis contig sequence #399.

XX Computer readable medium; Enterococcus faecalis; microbe; growth;
XX pathogenicity; vaccine; resistance; Enterococcal infection; commercial;
XX therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;
XX biochip technology; antibacterial; modulator of nucleic acid expression;
XX contig; ds.

XX Enterococcus faecalis.

OS US2002120116-A1.

XX 29-AUG-2002.

XX 04-MAY-1998; 98US-00070927.

XX 04-MAY-1998; 98US-00070927.

XX (KUNSCH) KUNSCH C A.

XX (DILLON) DILLON P J.

XX (BARASH) BARASH S.

XX Kunsch CA, Dillon PJ, Barash S;

XX WPI; 2002-750065/81.

XX Computer readable medium having recorded on it a Enterococcus faecalis
XX nucleotide sequence useful for detecting diseases related to Enterococcus
XX infections in animals.

XX Claim 1; Page; 119pp; English.

XX The present invention relates to a new computer readable medium with an
XX Enterococcus faecalis nucleotide sequence. The invention is useful to
XX diagnose the presence of E. faecalis in a sample or determining the
XX presence of a specific microbe in a sample. The invention is also useful
XX for modulating the growth or pathogenicity of E. faecalis, in a vaccine to
XX confer resistance to Enterococcal infection, for commercial, therapeutic
XX and industrial purposes, and for fermenting a particular sugar source or
XX to produce a particular metabolite. The invention is useful for detecting
XX diseases related to Enterococcus infections in animals, and for detecting
XX E. faecalis using biochip technology. The present nucleic acid sequence
XX represents an Enterococcus faecalis contig DNA sequence of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at <http://seqdata.uspto.gov>

XX Sequence 32768 BP; 10797 A; 5960 C; 7083 G; 8921 T; 0 U; 7 Other;

Alignment Scores:
Pred. No.: 7.35e+04 Length: 32768
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 90.2% Indels: 0

DB: 6 Gaps: 0
US-10-774-176-8 (1-9) x ABS99131 (1-32768)

Qy 1 AlallePheLeuValleuTyrIeu 9
Db 30913 GCCATTTTCTCTCGTCTCTTTTG 30939

RESULT 48
AAD31364
ID AAD31364 standard; DNA; 92139 BP.

XX AAD31364;
XX 31-MAY-2002 (first entry)

XX 92Kb gene fragment in human chromosome 17 at 17q21.

XX Human; Van Buchem's disease; genomic deletion; craniotubular hypertosis;
XX autosomal recessive disorder; chromosome 17; chromosome 17q21,
XX bone dysplasia; 92Kb gene fragment; db.

XX Homo sapiens.

XX Key Location/Qualifiers

XX misc_feature 5799..57515

XX FT /tag= a

XX FT /note= "This region is deleted in individuals afflicted

XX or carriers of Van Buchem's disease"

XX WO200210455-A2.

XX 07-FEB-2002.

XX 30-JUL-2001; 2001WO-US023968.

XX 28-JUL-2000; 2000US-0221855P.

XX 06-JUL-2001; 2001US-0303386P.

XX (CELL-) CELLTECH R & D INC.

XX (STRA/) STRAHLING HAMPTON K.

XX Brunkow ME, Proll S, Paepfer B;

XX WPI; 2002-227089/28.

XX Methods for identifying subjects who are afflicted with or carriers of
XX diseases associated with genomic deletion(s), e.g. Van Buchem's disease,
XX by determining the presence of a deletion in the 92 kb region of human
XX chromosome 17 at 17q21.

XX Claim 14; Page 45-72; 109pp; English.

XX The present invention relates to methods for distinguishing between
XX individuals homozygous for and therefore afflicted with Van Buchem's
XX disease, individuals heterozygous for and therefore carriers of Van
XX Buchem's disease and individuals who are not afflicted with Van Buchem's
XX disease comprise identifying a large genomic deletion in chromosome 17 at
XX 17q21. The method is useful for identifying individuals who are afflicted
XX with or carriers of diseases associated with one or more genomic
XX deletion, particularly Van Buchem's disease, which is a rare autosomal
XX recessive disorder that results in a bone dysplasia referred to as
XX craniotubular hypertosis. The present sequence is a 92Kb gene fragment in
XX human chromosome 17 at 17q21

XX Sequence 92139 BP; 23017 A; 22243 C; 23264 G; 23612 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 2.13e+05 Length: 92139
Score: 37.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.2% Indels: 0

DR WPI; 2003-328604/31.

XX
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.

XX
XX
PS Claim 1; SEQ ID NO 238; Opp; English.

XX
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published

XX
SQ Sequence 127369 BP; 37839 A; 19829 C; 21590 G; 38607 T; 0 U; 9504 Other;

Alignment Scores:			
Pred. No.:	2.96e+05	Length:	127369
Score:	37.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	90.2%	Indels:	0
DB:	11	Gaps:	0

US-10-774-176-8 (1-9) x ACN44006 (1-127369)

Qy	2	IlePheLeuLeuValIleuTyIreu	9
Db	102535	ATATTCCTCTAGTCTCTATTTA	102512

Search completed: April 25, 2006, 12:34:36
Job time : 340.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignments)
171.290 Million cell updates/sec

Title: US-10-774-176-8
Perfect score: 41
Sequence: 1 AIFLLVLYL 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB spool/US10774176/runat_24042006_165114_19197/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUPFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=1000
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes04
-USER=US10774176 @CCN 1 1 6765 @runat_24042006_165114_19197 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

GenEmbl:*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	290	6	CQ687716 Sequence
2	41	100.0	475	6	CQ920916 Sequence
3	41	100.0	901	6	BD249733 Polypepti

4	41	100.0	901	6	AX025013	AX025013 Sequence
5	41	100.0	901	6	AX316088	AX316088 Sequence
6	41	100.0	927	6	AX829164	AX829164 Sequence
7	41	100.0	1260	6	AX467373	AX467373 Sequence
8	41	100.0	1260	6	AX821533	AX821533 Sequence
9	41	100.0	1260	6	AX821548	AX821548 Sequence
10	41	100.0	1263	6	BD249731	BD249731 Polypepti
11	41	100.0	1263	6	AX025011	AX025011 Sequence
12	41	100.0	1263	6	AX149553	AX149553 Sequence
13	41	100.0	1263	6	AX316086	AX316086 Sequence
14	41	100.0	1263	6	AX467371	AX467371 Sequence
15	41	100.0	1281	6	BD249732	BD249732 Polypepti
16	41	100.0	1281	6	AX025012	AX025012 Sequence
17	41	100.0	1281	6	AX316087	AX316087 Sequence
18	41	100.0	2053	6	CQ731678	CQ731678 Sequence
19	41	100.0	2053	8	HS5740A	Z29083 Homo sapien
20	41	100.0	2333	9	AF063939	AF063939 Rattus no
21	41	100.0	2339	6	BD127282	BD127282 Primer fo
22	41	100.0	2359	6	CQ782724	CQ782724 Sequence
23	41	100.0	2359	8	AK074786	AK074786 Homo sapi
24	41	100.0	2361	6	BD127283	BD127283 Primer fo
25	41	100.0	2361	6	CQ782726	CQ782726 Sequence
26	41	100.0	2361	6	AX961916	AX961916 Sequence
27	41	100.0	2361	8	AK074790	AK074790 Homo sapi
28	41	100.0	2361	9	BC087011	BC087011 Rattus no
29	41	100.0	2379	8	BC037161	BC037161 Homo sapi
30	41	100.0	2423	9	BC058198	BC058198 Mus muscu
31	41	100.0	2557	6	AX961912	AX961912 Sequence
32	41	100.0	2557	6	AX961914	AX961914 Sequence
33	41	100.0	2714	8	AB168308	AB168308 Macaca fa
34	41	100.0	5551	8	HS4012159	AB012159 Homo sapi
35	41	100.0	7942	9	MMU012160	AU012160 Mus muscu
36	41	100.0	110000	15	AP008208	Continuation (358
37	41	100.0	120743	15	AP004082	AP004082 Oryza sat
38	41	100.0	121909	8	HSJ492P14	AL121977 Human DNA
39	41	100.0	129010	8	AL589666	AL589666 Human DNA
40	41	100.0	167046	9	AC158516	AC158516 Mus muscu
41	41	100.0	208985	14	AC155741	AC155741 Bos tauru
42	41	100.0	210237	14	AC128294	AC128294 Rattus no
43	41	100.0	218269	14	AC156733	AC156733 Bos tauru
44	41	100.0	239076	14	AC106962	AC106962 Rattus no
45	40	97.6	397	3	UBA518259	AJ518259 Unidentif
46	40	97.6	12448	1	AB001270	AB001270 Treponema
47	40	97.6	110000	1	AY596297	Continuation (14 o
48	40	97.6	164489	9	AC121542	AC121542 Mus muscu
49	40	97.6	212987	14	AC163932	AC163932 Bos tauru
50	39	95.1	63024	14	AC101062	AC101062 Mus muscu
51	39	95.1	108124	8	AL513206	AL513206 Human DNA
52	39	95.1	110000	14	CT009751	Continuation (2 of
53	39	95.1	110000	14	CT009753	Continuation (4 of
54	39	95.1	110000	15	CR382137	Continuation (9 of
55	39	95.1	130191	14	AC162138	AC162138 Loxodonta
56	39	95.1	156815	9	AC125165	AC125165 Mus muscu
57	39	95.1	158654	9	AC138110	AC138110 Mus muscu
58	39	95.1	164461	8	AC083806	AC083806 Homo sapi
59	39	95.1	171215	8	AC099618	AC099618 Mus muscu
60	39	95.1	207701	9	AC154196	AC154196 Mus muscu
61	39	95.1	220403	14	CT010571	CT010571 Mus muscu
62	39	95.1	224956	14	AC096811	AC096811 Rattus no
63	39	95.1	246036	14	AC097816	AC097816 Rattus no
64	39	95.1	264977	14	AC097564	AC097564 Rattus no
65	39	95.1	283967	14	AC105589	AC105589 Rattus no
66	38	92.7	33297	8	AF293358	AF293358 Homo sapi
67	38	92.7	86296	8	AL590224	AL590224 Human DNA
68	38	92.7	86973	8	AC011394	AC011394 Homo sapi
69	38	92.7	92648	14	AP007853	Continuation (4 of
70	38	92.7	97268	5	CR933813	CR933813 Zebrafish
71	38	92.7	101447	14	AC152432	AC152432 Bos tauru
72	38	92.7	110000	1	AB017220	Continuation (17 o
73	38	92.7	110000	14	AL928996	AL928996 Mus muscu
74	38	92.7	110000	15	AP008213	Continuation (60 o
75	38	92.7	120823	15	AP003748	AP003748 Oryza sat

77	38	92.7	126681	9	AL645636	AL645636 Mouse DNA	150	37	90.2	90020	15	ATT24H18	AL353013 Arabidops
c 78	38	92.7	152249	14	CR450747	CR450747 Danio rer	151	37	90.2	91587	14	AC112924	AC112924 Mus muscu
c 79	38	92.7	154078	8	AC123767	AC123767 Homo sapi	152	37	90.2	92139	6	AR242855	AR242855 Sequence
c 80	38	92.7	157205	15	AP005098	AP005098 Oryza sat	153	37	90.2	92139	6	AX384907	AX384907 Sequence
81	38	92.7	160280	14	AC146628	AC146628 Papio anu	c 154	37	90.2	93790	8	AF397423	AF397423 Homo sapi
82	38	92.7	161424	14	AC018367	AC018367 Homo sapi	155	37	90.2	96543	9	BX005138	BX005138 Mouse DNA
83	38	92.7	162128	9	AP003156	AP003156 Mus muscu	c 156	37	90.2	104281	15	AC129092	AC129092 Medicago
84	38	92.7	172357	14	AP001808	AP001808 Homo sapi	c 157	37	90.2	110000	1	CR522870	Continuation (14 o
85	38	92.7	172655	14	AC146627	AC146627 Papio anu	158	37	90.2	110000	6	BD426631	Continuation (17 o
86	38	92.7	173856	14	AC007526	AC007526 Homo sapi	159	37	90.2	110000	6	AR627451	Continuation (17 o
87	38	92.7	179084	14	AC157621	AC157621 Papio anu	160	37	90.2	110000	6	CR632719	Continuation (17 o
c 88	38	92.7	180261	9	AC154546	AC154546 Mus muscu	161	37	90.2	110000	14	CR940853	Continuation (3 of
c 89	38	92.7	181552	8	AC012417	AC012417 Homo sapi	162	37	90.2	110000	14	CR926457	Continuation (2 of
c 90	38	92.7	193722	14	AC162641	AC162641 Bos tauru	c 163	37	90.2	110000	15	AP008218	Continuation (147
91	38	92.7	194260	9	AP003158	AP003158 Mus muscu	c 164	37	90.2	110000	15	AP008210	Continuation (331
92	38	92.7	195331	14	AC068174	AC068174 Homo sapi	c 165	37	90.2	119328	15	AC130801	AC130801 Medicago
c 93	38	92.7	198714	8	CNS05YDM	AL357172 Homo chr	166	37	90.2	120185	15	AC007138	AC007138 Arabidops
c 94	38	92.7	202212	8	CNS01RGP	AL357957 Human chr	167	37	90.2	120913	9	AL844545	AL844545 Mouse DNA
c 95	38	92.7	202490	14	AP002426	AP002426 Homo sapi	168	37	90.2	122336	8	AC003075	AC003075 Homo sapi
96	38	92.7	204556	14	AC153750	AC153750 Callicebu	169	37	90.2	125898	5	CR392366	CR392366 Zebrafish
97	38	92.7	215521	14	AC126103	AC126103 Rattus no	170	37	90.2	128873	9	AC133171	AC133171 Mus muscu
c 98	38	92.7	217766	9	AC087558	AC087558 Mus muscu	171	37	90.2	131366	8	AL354701	AL354701 Human DNA
c 99	38	92.7	227236	14	AC136709	AC136709 Mus muscu	172	37	90.2	131673	15	CNS08CCX	AL954159 Oryza sat
100	38	92.7	233337	14	AC109704	AC109704 Rattus no	c 173	37	90.2	131804	14	AC161722	AC161722 Loxodonta
c 101	38	92.7	234302	14	AC098958	AC098958 Rattus no	c 174	37	90.2	136937	14	AC139627	AC139627 Takifugu
c 102	38	92.7	245388	14	AC133717	AC133717 Rattus no	c 175	37	90.2	138419	9	AC111060	AC111060 Mus muscu
103	38	92.7	263028	14	AC164222	AC164222 Bos tauru	c 176	37	90.2	138981	8	AC141002	AC141002 Homo sapi
104	38	92.7	269596	14	AC164050	AC164050 Bos tauru	c 177	37	90.2	140626	15	OSJN00057	AL606635 Oryza sat
c 105	38	92.7	269596	14	AC164050	AC164050 Bos tauru	178	37	90.2	141307	9	AC135509	AC135509 Mus muscu
c 106	37	90.2	154	10	BV167253	BV167253 sqm4262	179	37	90.2	141515	14	CR846092	CR846092 Danio rer
107	37	90.2	204	6	AR396453	AR396453 Sequence	c 180	37	90.2	143413	14	AC025979	AC025979 Homo sapi
c 108	37	90.2	421	10	HSP3A2B9	Z94208 H. sapiens f	c 181	37	90.2	143947	8	AC078833	AC078833 Homo sapi
c 109	37	90.2	634	10	BV041847	BV041847 S212P6914	c 182	37	90.2	145013	5	CR382283	CR382283 Zebrafish
c 110	37	90.2	644	10	BV441077	BV441077 S237P684R	183	37	90.2	145676	8	AL161786	AL161786 Human DNA
c 111	37	90.2	945	2	AF337019	AF337019 Aphidius	c 184	37	90.2	146137	14	AC148795	AC148795 Otollemur
c 112	37	90.2	1058	6	C0578805	C0578805 Sequence	c 185	37	90.2	146311	8	AC147032	AC147032 Pan trogl
113	37	90.2	1160	5	CR523763	CR523763 Gallus ga	c 186	37	90.2	149189	5	AL845366	AL845366 Zebrafish
114	37	90.2	1219	9	RATBGASILI	M73985 Rat liver-s	c 187	37	90.2	151024	14	AC016855	AC016855 Homo sapi
115	37	90.2	1228	15	ATH243456	AJ243456 Arabidops	c 188	37	90.2	151862	14	BX936339	BX936339 Danio rer
116	37	90.2	1293	5	BC066508	BC066508 Danio rer	c 189	37	90.2	152896	14	AX936339	AX936339 Canis fam
117	37	90.2	1422	15	AY084857	AY084857 Arabidops	c 190	37	90.2	155867	8	HS404H4	AL031661 Human DNA
118	37	90.2	1510	15	DQ013105	DQ013105 Phallocep	c 191	37	90.2	156139	14	AP001174	AP001174 Homo sapi
119	37	90.2	2025	5	BC055653	BC055653 Danio rer	c 192	37	90.2	156420	14	AC158215	AC158215 Mus muscu
120	37	90.2	2183	5	CR855786	CR855786 Xenopus t	c 193	37	90.2	157691	8	AC009051	AC009051 Homo sapi
121	37	90.2	2210	15	SCSECEG1G	X62340 S. cerevisia	c 194	37	90.2	158261	15	CNS08C9S	AL772414 Oryza sat
122	37	90.2	2265	5	BC065654	BC065654 Danio rer	c 195	37	90.2	158657	2	AY800246	AY800246 Tribolium
123	37	90.2	3298	6	C0578804	C0578804 Sequence	c 196	37	90.2	159264	14	AC092726	AC092726 Homo sapi
124	37	90.2	5888	2	AF230482	AF230482 Hydra vul	197	37	90.2	159867	8	AL450303	AL450303 Human DNA
125	37	90.2	6224	6	AX281313	AX281313 Sequence	c 198	37	90.2	160001	2	AC113619	AC113619 Drosophil
126	37	90.2	6224	6	AX346210	AX346210 Sequence	c 199	37	90.2	160724	5	BX005101	BX005101 Zebrafish
c 127	37	90.2	8096	2	DSRXDEGN	X98371 D. subobscur	c 200	37	90.2	160872	14	AC053534	AC053534 Homo sapi
c 128	37	90.2	9515	1	AF486554	AF486554 Campyloba	c 201	37	90.2	160890	8	AC146413	AC146413 Pan trogl
c 129	37	90.2	9561	1	AF486556	AF486556 Campyloba	c 202	37	90.2	161705	8	AC158799	AC158799 Mus muscu
c 130	37	90.2	9568	1	AF486552	AF486552 Campyloba	c 203	37	90.2	162351	8	AC123303	AC123303 Homo sapi
131	37	90.2	10952	1	U32835	U32835 Haemophilus	c 204	37	90.2	162676	9	AL138958	AL138958 Human DNA
c 132	37	90.2	22761	15	Y8CL3502	U91904 Saccharomyc	c 205	37	90.2	162973	8	AL138958	AL138958 Human DNA
c 133	37	90.2	26552	14	AC132682	AC132682 Homo sapi	c 206	37	90.2	163198	14	AL161634	AL161634 Homo sapi
134	37	90.2	27598	2	CRP55A11	Z72511 Caenorhabdi	c 207	37	90.2	163352	8	AC009225	AC009225 Homo sapi
135	37	90.2	37368	6	BD193805	BD193805 Enterococ	c 208	37	90.2	163690	14	AC140908	AC140908 Homo sapi
c 136	37	90.2	33318	14	AC131089	AC131089 Homo sapi	c 209	37	90.2	164238	8	AC106745	AC106745 Homo sapi
c 137	37	90.2	33901	14	AC127087	AC127087 Homo sapi	c 210	37	90.2	165699	8	AL513023	AL513023 Human DNA
c 138	37	90.2	43586	14	OSIG00036	AL732335 Oryza sat	c 211	37	90.2	167310	9	AC105169	AC105169 Mus muscu
c 139	37	90.2	50630	8	AC107054	AC107054 Homo sapi	c 212	37	90.2	167853	14	AP001798	AP001798 Homo sapi
c 140	37	90.2	63704	14	AC087469	AC087469 Homo sapi	c 213	37	90.2	168752	9	AC138118	AC138118 Mus muscu
c 141	37	90.2	64079	8	AL137841	AL137841 Human DNA	c 214	37	90.2	169717	14	AC104856	AC104856 Homo sapi
c 142	37	90.2	66145	15	H0609A12	AL512544 Oryza sat	c 215	37	90.2	169964	8	AC055710	AC055710 Mus muscu
c 143	37	90.2	66586	14	AC090183	AC090183 Homo sapi	c 216	37	90.2	170232	8	AL135903	AL135903 Human DNA
c 144	37	90.2	66586	14	AC090183	AC090183 Homo sapi	c 217	37	90.2	171823	9	AC126800	AC126800 Mus muscu
c 145	37	90.2	71197	14	AC091673	AC091673 Homo sapi	c 218	37	90.2	172692	5	BX005380	BX005380 Zebrafish
c 146	37	90.2	71516	14	AC014220	AC014220 Drosophil	c 219	37	90.2	173736	9	AC139569	AC139569 Mus muscu
c 147	37	90.2	80117	8	AC055813	AC055813 Homo sapi	c 220	37	90.2	173821	8	AC020599	AC020599 Homo sapi
c 148	37	90.2	83885	14	AC074301	AC074301 Homo sapi	c 221	37	90.2	174176	8	AC023263	AC023263 Homo sapi
149	37	90.2	87503	15	ATT8H10	AL133248 Arabidops	c 222	37	90.2	174337	14	AC012232	AC012232 Homo sapi

C 223	37	90.2	175111	8	AC006992	AC006992 Homo sapi	C 296	90.2	240406	9	AC103939	AC103939 Mus muscu
C 224	37	90.2	176545	1	AE016957	AE016957 Enterococ	C 297	90.2	241248	14	AC096799	AC096799 Rattus no
C 225	37	90.2	177176	14	AC139877	AC139877 Didelphis	C 298	90.2	242896	14	CR853288	CR853288 Danio rer
C 226	37	90.2	178033	8	AC107025	AC107025 Homo sapi	C 299	90.2	243833	14	AC097386	AC097386 Rattus no
C 227	37	90.2	178415	14	AC023547	AC023547 Homo sapi	C 300	90.2	246017	14	AC119760	AC119760 Rattus no
C 228	37	90.2	179227	8	AC069223	AC069223 Homo sapi	C 301	90.2	247220	14	AC073807	AC073807 Mus muscu
C 229	37	90.2	180262	8	AC009161	AC009161 Homo sapi	C 302	90.2	247892	9	AC113316	AC113316 Mus muscu
C 230	37	90.2	180392	8	AC025945	AC025945 Homo sapi	C 303	90.2	248036	14	AC155043	AC155043 Bos tauru
C 231	37	90.2	180838	8	AC068763	AC068763 Homo sapi	C 304	90.2	248594	14	AC106590	AC106590 Rattus no
C 232	37	90.2	181663	5	CR354436	CR354436 Zebrafish	C 305	90.2	250038	14	AC131878	AC131878 Rattus no
C 233	37	90.2	182661	9	AC121602	AC121602 Mus muscu	C 306	90.2	250713	14	AC103517	AC103517 Rattus no
C 234	37	90.2	183502	14	AC150601	AC150601 Callithiri	C 307	90.2	251012	14	AC157404	AC157404 Bos tauru
C 235	37	90.2	184635	8	AC090239	AC090239 Homo sapi	C 308	90.2	253429	14	AC095746	AC095746 Rattus no
C 236	37	90.2	184857	14	AP002503	AP002503 Homo sapi	C 309	90.2	257152	14	AC119134	AC119134 Rattus no
C 237	37	90.2	185237	8	AC011124	AC011124 Homo sapi	C 310	90.2	257237	14	CR926458	CR926458 Danio rer
C 238	37	90.2	185966	8	AC155320	AC155320 Pan trogl	C 311	90.2	261137	14	AC152706	AC152706 Bos tauru
C 239	37	90.2	186044	14	AC087674	AC087674 Homo sapi	C 312	90.2	264051	14	AC009050	AC009050 Homo sapi
C 240	37	90.2	186101	14	AC161644	AC161644 Bos tauru	C 313	90.2	267396	14	CR352293	CR352293 Danio rer
C 241	37	90.2	187030	8	AC131097	AC131097 Homo sapi	C 314	90.2	276204	14	AC154998	AC154998 Bos tauru
C 242	37	90.2	187389	14	AC121052	AC121052 Rattus no	C 315	90.2	300162	14	AC149441	AC149441 Meleagris
C 243	37	90.2	187772	9	AC123947	AC123947 Mus muscu	C 316	90.2	305376	2	AB003437	AB003437 Drosophil
C 244	37	90.2	188486	14	AP001796	AP001796 Homo sapi	C 317	90.2	306464	14	AC068708	AC068708 Homo sapi
C 245	37	90.2	189976	2	AC023710	AC023710 Drosophil	C 318	90.2	349122	1	BX569694	BX569694 Synchoco
C 246	37	90.2	190394	8	AL157834	AL157834 Human DNA	C 319	90.2	349122	1	BX569694	BX569694 Synchoco
C 247	37	90.2	191353	14	AC090320	AC090320 Homo sapi	C 320	90.2	349122	1	BX569694	BX569694 Synchoco
C 248	37	90.2	192244	14	AP001382	AP001382 Homo sapi	C 321	90.2	349122	1	BX569694	BX569694 Synchoco
C 249	37	90.2	192869	9	AC139755	AC139755 Mus muscu	C 322	90.2	349122	1	BX569694	BX569694 Synchoco
C 250	37	90.2	193207	9	AC121968	AC121968 Homo sapi	C 323	90.2	349122	1	BX569694	BX569694 Synchoco
C 251	37	90.2	193569	9	AC100795	AC100795 Homo sapi	C 324	90.2	349122	1	BX569694	BX569694 Synchoco
C 252	37	90.2	194431	8	AL157888	AL157888 Human DNA	C 325	90.2	349122	1	BX569694	BX569694 Synchoco
C 253	37	90.2	195373	14	AC107849	AC107849 Mus muscu	C 326	90.2	349122	1	BX569694	BX569694 Synchoco
C 254	37	90.2	195485	8	AL591846	AL591846 Human DNA	C 327	90.2	349122	1	BX569694	BX569694 Synchoco
C 255	37	90.2	195923	8	AC022489	AC022489 Homo sapi	C 328	90.2	349122	1	BX569694	BX569694 Synchoco
C 256	37	90.2	196036	14	AC145239	AC145239 Pongo pyg	C 329	90.2	349122	1	BX569694	BX569694 Synchoco
C 257	37	90.2	198220	15	ATCHRIV5	ATCHRIV5 Arabidops	C 330	90.2	349122	1	BX569694	BX569694 Synchoco
C 258	37	90.2	198727	9	AC115062	AC115062 Mus muscu	C 331	90.2	349122	1	BX569694	BX569694 Synchoco
C 259	37	90.2	199740	5	BX449299	BX449299 Zebrafish	C 332	90.2	349122	1	BX569694	BX569694 Synchoco
C 260	37	90.2	200959	14	CR847547	CR847547 Danio rer	C 333	90.2	349122	1	BX569694	BX569694 Synchoco
C 261	37	90.2	201416	14	AC130437	AC130437 Homo sapi	C 334	90.2	349122	1	BX569694	BX569694 Synchoco
C 262	37	90.2	202907	14	AC106401	AC106401 Rattus no	C 335	90.2	349122	1	BX569694	BX569694 Synchoco
C 263	37	90.2	204478	14	AC130948	AC130948 Rattus no	C 336	90.2	349122	1	BX569694	BX569694 Synchoco
C 264	37	90.2	206466	14	AC109663	AC109663 Rattus no	C 337	90.2	349122	1	BX569694	BX569694 Synchoco
C 265	37	90.2	208077	14	AC129811	AC129811 Rattus no	C 338	90.2	349122	1	BX569694	BX569694 Synchoco
C 266	37	90.2	209444	14	AC079619	AC079619 Homo sapi	C 339	90.2	349122	1	BX569694	BX569694 Synchoco
C 267	37	90.2	211624	9	AC105631	AC105631 Rattus no	C 340	90.2	349122	1	BX569694	BX569694 Synchoco
C 268	37	90.2	211817	14	AC128557	AC128557 Rattus no	C 341	90.2	349122	1	BX569694	BX569694 Synchoco
C 269	37	90.2	214393	9	BX537302	BX537302 Mouse DNA	C 342	90.2	349122	1	BX569694	BX569694 Synchoco
C 270	37	90.2	215437	9	AC129336	AC129336 Mus muscu	C 343	90.2	349122	1	BX569694	BX569694 Synchoco
C 271	37	90.2	215861	9	AC140313	AC140313 Mus muscu	C 344	90.2	349122	1	BX569694	BX569694 Synchoco
C 272	37	90.2	218647	9	AC152397	AC152397 Mus muscu	C 345	90.2	349122	1	BX569694	BX569694 Synchoco
C 273	37	90.2	218708	14	BX897744	BX897744 Danio rer	C 346	90.2	349122	1	BX569694	BX569694 Synchoco
C 274	37	90.2	218859	14	AC022912	AC022912 Homo sapi	C 347	90.2	349122	1	BX569694	BX569694 Synchoco
C 275	37	90.2	219015	14	AC156419	AC156419 Bos tauru	C 348	90.2	349122	1	BX569694	BX569694 Synchoco
C 276	37	90.2	220296	14	AC110699	AC110699 Rattus no	C 349	90.2	349122	1	BX569694	BX569694 Synchoco
C 277	37	90.2	222518	14	AC096499	AC096499 Rattus no	C 350	90.2	349122	1	BX569694	BX569694 Synchoco
C 278	37	90.2	225585	14	AC113236	AC113236 Canis fam	C 351	90.2	349122	1	BX569694	BX569694 Synchoco
C 279	37	90.2	226448	14	AC097284	AC097284 Rattus no	C 352	90.2	349122	1	BX569694	BX569694 Synchoco
C 280	37	90.2	227279	14	AC120785	AC120785 Mus muscu	C 353	90.2	349122	1	BX569694	BX569694 Synchoco
C 281	37	90.2	228745	14	AC128310	AC128310 Rattus no	C 354	90.2	349122	1	BX569694	BX569694 Synchoco
C 282	37	90.2	228802	14	AC023717	AC023717 Drosophil	C 355	90.2	349122	1	BX569694	BX569694 Synchoco
C 283	37	90.2	229048	14	AC156182	AC156182 Bos tauru	C 356	90.2	349122	1	BX569694	BX569694 Synchoco
C 284	37	90.2	230475	14	CR538727	CR538727 Danio rer	C 357	90.2	349122	1	BX569694	BX569694 Synchoco
C 285	37	90.2	230945	9	AC122898	AC122898 Mus muscu	C 358	90.2	349122	1	BX569694	BX569694 Synchoco
C 286	37	90.2	231623	5	BX255933	BX255933 Zebrafish	C 359	90.2	349122	1	BX569694	BX569694 Synchoco
C 287	37	90.2	231868	14	AC118124	AC118124 Rattus no	C 360	90.2	349122	1	BX569694	BX569694 Synchoco
C 288	37	90.2	234354	9	AC116406	AC116406 Mus muscu	C 361	90.2	349122	1	BX569694	BX569694 Synchoco
C 289	37	90.2	234405	14	AC095507	AC095507 Rattus no	C 362	90.2	349122	1	BX569694	BX569694 Synchoco
C 290	37	90.2	235286	8	AC136285	AC136285 Homo sapi	C 363	90.2	349122	1	BX569694	BX569694 Synchoco
C 291	37	90.2	235419	14	AC099361	AC099361 Rattus no	C 364	90.2	349122	1	BX569694	BX569694 Synchoco
C 292	37	90.2	236262	14	AC118080	AC118080 Rattus no	C 365	90.2	349122	1	BX569694	BX569694 Synchoco
C 293	37	90.2	237547	14	BX539339	BX539339 Mus muscu	C 366	90.2	349122	1	BX569694	BX569694 Synchoco
C 294	37	90.2	239057	8	AP001578	AP001578 Homo sapi	C 367	90.2	349122	1	BX569694	BX569694 Synchoco
C 295	37	90.2	239847	14	AC128434	AC128434 Rattus no	C 368	90.2	349122	1	BX569694	BX569694 Synchoco

369	36	87.8	8457	13	AAVSPHERE	M77182 Amsacta ent	c 442	36	87.8	110000	1	BA000004_14	Continuation (15 o
370	36	87.8	9963	1	AE008421	AE008421 Streptoco	443	36	87.8	110000	1	BA000028_33	Continuation (34 o
371	36	87.8	11211	1	AE010213	AE010213 Pyrococu	444	36	87.8	110000	1	BA000043_33	Continuation (34 o
372	36	87.8	13305	8	AL449163	AL449163 Human DNA	445	36	87.8	110000	1	CP000002_07	Continuation (8 of
373	36	87.8	15212	1	U32779	U32779 Haemophilus	446	36	87.8	110000	1	CP000057_10	Continuation (11 o
374	36	87.8	16457	14	AC101662_4	Continuation (5 of	c 447	36	87.8	110000	2	AC116984_4	Continuation (5 of
375	36	87.8	21961	2	CEP14F4	AL021446 Caenorhab	448	36	87.8	110000	6	BD426631_10	Continuation (11 o
376	36	87.8	24830	9	AF081501	AF081501 Mus muscu	449	36	87.8	110000	6	AR274513_10	Continuation (11 o
377	36	87.8	34173	2	AF047655	AF047655 Caenorhab	450	36	87.8	110000	6	AR632719_10	Continuation (11 o
378	36	87.8	35948	8	AC128659	AC128659 Homo sapi	c 451	36	87.8	110000	14	AC106339_0	AC106339 Rattus no
379	36	87.8	37494	2	AF078785	AF078785 Caenorhab	c 452	36	87.8	110000	14	AC106339_1	Continuation (2 of
380	36	87.8	38189	14	AC165052	AC165052 Phakopseor	453	36	87.8	110000	14	AC107137_1	Continuation (2 of
381	36	87.8	38490	2	AC132213	AC132213 Drosophil	454	36	87.8	110000	14	AC107137_2	Continuation (3 of
382	36	87.8	40889	2	AC158490	AC158490 Capitella	c 455	36	87.8	110000	14	AC107144_2	Continuation (3 of
383	36	87.8	41700	14	AP000634	AP000634 Homo sapi	456	36	87.8	110000	14	AC101048_1	Continuation (2 of
384	36	87.8	43129	14	AC131379	AC131379 Strongylo	c 457	36	87.8	110000	14	AC110929_4	Continuation (5 of
385	36	87.8	44725	8	AF036405	AF036405 Homo sapi	458	36	87.8	110000	14	AC111224_1	Continuation (2 of
386	36	87.8	45045	8	AL359954	AL359954 Human DNA	c 459	36	87.8	110000	14	AC113216_1	Continuation (2 of
387	36	87.8	45373	8	AL158092	AL158092 Human DNA	460	36	87.8	110000	14	AC113216_4	Continuation (5 of
388	36	87.8	50000	6	AX392736	AX392736 Sequence	461	36	87.8	110000	14	AC118915_2	Continuation (3 of
389	36	87.8	50765	14	CR854841_3	Continuation (4 of	c 462	36	87.8	110000	14	AC127184_1	Continuation (2 of
390	36	87.8	51191	8	AC003951	AC003951 Homo sapi	c 463	36	87.8	110000	14	AC161380_3	Continuation (4 of
391	36	87.8	51410	8	AC092604	AC092604 Homo sapi	c 464	36	87.8	110000	14	AL390072_2	Continuation (3 of
392	36	87.8	52989	5	CR354400	CR354400 Zebrafish	c 465	36	87.8	110000	14	AL928996_0	AL928996 Mus muscu
393	36	87.8	61619	9	AC124785	AC124785 Rattus no	466	36	87.8	110000	14	AL928996_5	Continuation (6 of
394	36	87.8	6252	5	BX322549	BX322549 Zebrafish	c 467	36	87.8	110000	14	BX511044_1	BX511044 Homo sapi
395	36	87.8	62766	14	AC102400	AC102400 Mus muscu	c 468	36	87.8	110000	14	BX511044_2	Continuation (2 of
396	36	87.8	63325	8	AL353592	AL353592 Human DNA	c 469	36	87.8	110000	14	BX511044_3	Continuation (3 of
397	36	87.8	65017	14	AC115958	AC115958 Mus muscu	470	36	87.8	110000	14	CR382381_1	Continuation (2 of
398	36	87.8	66857	14	AC068239	AC068239 Homo sapi	471	36	87.8	110000	14	CR382381_2	Continuation (2 of
399	36	87.8	67085	14	AC104937	AC104937 Homo sapi	472	36	87.8	110000	14	CT005270_12	Continuation (13 o
400	36	87.8	67776	14	AC101220	AC101220 Mus muscu	c 473	36	87.8	110000	15	AP008215_149	Continuation (15 o
401	36	87.8	67911	8	AP001413	AP001413 Homo sapi	c 474	36	87.8	110000	15	CR382125_02	Continuation (3 of
402	36	87.8	67970	2	PFMAL1P3	AL031746 Plasmodiu	475	36	87.8	110000	15	AE017356_6	Continuation (7 of
403	36	87.8	69941	14	AC027815	AC027815 Homo sapi	c 476	36	87.8	110000	15	AP008207_043	Continuation (44 o
404	36	87.8	71120	14	AP005417	AP005417 Oryza sat	c 477	36	87.8	110000	15	AP008208_297	Continuation (298
405	36	87.8	72443	14	AC015973	AC015973 Drosophil	478	36	87.8	110000	15	AP008212_083	Continuation (84 o
406	36	87.8	72636	14	AC013201	AC013201 Drosophil	c 479	36	87.8	110000	15	AP008213_066	Continuation (67 o
407	36	87.8	72660	14	AC164775	AC164775 Bos tauru	c 480	36	87.8	110898	8	AL355973	AL355973 Human DNA
408	36	87.8	76066	8	AL162588	AL162588 Human DNA	c 481	36	87.8	112007	5	CR361545	CR361545 Zebrafish
409	36	87.8	77969	15	AP006347	AP006347 Lotu cor	482	36	87.8	112074	8	AP000012	AP000012 Homo sapi
410	36	87.8	79729	14	AC165503	AC165503 Bos tauru	c 483	36	87.8	112972	2	AC0006470	AC0006470 Drosophil
411	36	87.8	80734	9	AL929550	AL929550 Mouse DNA	484	36	87.8	114080	8	AC133123	AC133123 Homo sapi
412	36	87.8	81567	15	AY730338	AY730338 Solanum t	485	36	87.8	114287	15	AC157472	AC157472 Medicago
413	36	87.8	83373	15	AB017064	AB017064 Arabidops	486	36	87.8	114546	14	HCAC000382	HCAC000382 Homo sapi
414	36	87.8	83463	8	AL161644	AL161644 Human DNA	c 487	36	87.8	115508	14	AC160308	AC160308 Bos tauru
415	36	87.8	85237	15	AC139743	AC139743 Medicago	488	36	87.8	117864	15	AC139854	AC139854 Medicago
416	36	87.8	85599	15	AB023045	AB023045 Arabidops	489	36	87.8	118452	8	AL359677	AL359677 Human DNA
417	36	87.8	86723	8	AF401203	AF401203 Homo sapi	c 490	36	87.8	118662	8	AC108013	AC108013 Homo sapi
418	36	87.8	86958	8	AC004464	AC004464 Homo sapi	c 491	36	87.8	118767	9	BX000344	BX000344 Mouse DNA
419	36	87.8	89753	8	AY373585	AY373585 Homo sapi	c 492	36	87.8	120407	14	AP001967	AP001967 Homo sapi
420	36	87.8	91875	8	AL512286	AL512286 Human DNA	c 493	36	87.8	120722	14	AC162758	AC162758 Loxodonta
421	36	87.8	92127	14	AC111295_3	Continuation (4 of	494	36	87.8	122685	8	AC094023	AC094023 Homo sapi
422	36	87.8	95171	14	AC166424_2	AC166424 Bos tauru	495	36	87.8	122770	15	AC147007	AC147007 Medicago
423	36	87.8	95091	14	AC130040	AC130040 Rattus no	c 496	36	87.8	122913	8	AL139404	AL139404 Human DNA
424	36	87.8	97297	9	AL731677	AL731677 Homo sapi	c 497	36	87.8	123089	8	AC135731	AC135731 Homo sapi
425	36	87.8	98889	14	AL590863	AL590863 Homo sapi	498	36	87.8	123233	14	AC135625	AC135625 Homo sapi
426	36	87.8	100000	8	AP000154	AP000154 Homo sapi	c 499	36	87.8	123428	15	AP004098	AP004098 Oryza sat
427	36	87.8	101160	14	AP007356	AP007356 Lotu cor	c 500	36	87.8	124489	8	AC110807	AC110807 Homo sapi
428	36	87.8	101421	14	AC157537	AC157537 Medicago	c 501	36	87.8	124927	14	AC010475	AC010475 Homo sapi
429	36	87.8	102692	15	AC149288	AC149288 Solanum d	c 502	36	87.8	125439	8	AC010485	AC010485 Homo sapi
430	36	87.8	102965	14	AC146727	AC146727 Ootlemur	c 503	36	87.8	126135	8	AC107058	AC107058 Homo sapi
431	36	87.8	104140	8	AC008122	AC008122 Homo sapi	c 504	36	87.8	127105	14	AC011881	AC011881 Homo sapi
432	36	87.8	105604	8	AC084013	AC084013 Homo sapi	c 505	36	87.8	127374	9	AC107675	AC107675 Mus muscu
433	36	87.8	105898	8	AL590302	AL590302 Human DNA	506	36	87.8	127422	14	AC136294	AC136294 Homo sapi
434	36	87.8	105234	8	AY267352	AY267352 Homo sapi	507	36	87.8	127682	8	AL137861	AL137861 Human DNA
435	36	87.8	107328	8	AL160057	AL160057 Human DNA	508	36	87.8	129889	8	CR381670	CR381670 Human DNA
436	36	87.8	107641	9	AF125313	AF125313 Mus muscu	509	36	87.8	129921	15	AC146568	AC146568 Medicago
437	36	87.8	107816	15	AC016447	AC016447 Arabidops	510	36	87.8	130226	15	AC162299	AC162299 Mus muscu
438	36	87.8	109997	14	AC164464	AC164464 Bos tauru	c 511	36	87.8	130660	15	AC144564	AC144564 Medicago
439	36	87.8	110000	1	CP000099_08	Continuation (9 of	512	36	87.8	130726	8	AL365265	AL365265 Human DNA
440	36	87.8	110000	1	CP000099_16	Continuation (17 o	513	36	87.8	131056	8	CR381653	CR381653 Human DNA
441	36	87.8	110000	1	AE017333_07	Continuation (8 of	514	36	87.8	132329	14	AC161034	AC161034 Medicago

C 515	36	87.8	132490	9	AC116841	AC116841 Mus muscu	588	36	87.8	165504	9	AC140352	AC140352 Mus muscu
C 516	36	87.8	133787	8	AC068537	AC068537 Homo sapi	589	36	87.8	165549	8	AL355498	AL355498 Human DNA
C 517	36	87.8	133948	14	BX322233	BX322233 Homo sapi	590	36	87.8	165731	8	BX640538	BX640538 Human DNA
C 518	36	87.8	133961	8	AC161732	AC161732 Loxodonta	591	36	87.8	165796	9	AL124115	AL124115 Mus muscu
C 519	36	87.8	133572	14	AC004998	AC004998 Homo sapi	C 592	36	87.8	165880	8	AL136372	AL136372 Human DNA
C 520	36	87.8	137174	15	AP002484	AP002484 Oryza sat	C 593	36	87.8	165903	9	AP007208	AP007208 Mus muscu
C 521	36	87.8	137809	14	AC158977	AC158977 Mus muscu	C 594	36	87.8	165943	9	AP007208	AP007208 Mus muscu
C 522	36	87.8	138219	14	AC156347	AC156347 Otlemur	C 595	36	87.8	165948	8	AC018616	AC018616 Homo sapi
C 523	36	87.8	140675	5	AC156244	AC156244 Xenopus t	C 596	36	87.8	166118	8	AL355495	AL355495 Human DNA
C 524	36	87.8	141036	14	AC023807	AC023807 Mus muscu	597	36	87.8	166308	5	CR513784	CR513784 Zebrafish
C 525	36	87.8	141983	15	AP003047	AP003047 Oryza sat	598	36	87.8	166452	14	CR382332	CR382332 Homo sapi
C 526	36	87.8	142500	4	AC164543	AC164543 Monodelph	C 599	36	87.8	166484	14	AC026830	AC026830 Homo sapi
C 527	36	87.8	142616	14	AC067789	AC067789 Homo sapi	C 600	36	87.8	166523	14	AC022646	AC022646 Homo sapi
C 528	36	87.8	143003	8	AC159541	AC159541 Homo sapi	601	36	87.8	166555	9	AC124366	AC124366 Mus muscu
C 529	36	87.8	143116	14	AC023966	AC023966 Homo sapi	C 602	36	87.8	166704	14	AP001559	AP001559 Homo sapi
C 530	36	87.8	144201	8	HS230119	Z93942 Human DNA s	C 603	36	87.8	166965	14	AC122712	AC122712 Homo sapi
C 531	36	87.8	144854	9	AL670758	AL670758 Mouse DNA	C 604	36	87.8	167426	14	AC091085	AC091085 Homo sapi
C 532	36	87.8	145255	8	AP002372	AP002372 Homo sapi	C 605	36	87.8	167484	8	AC009446	AC009446 Homo sapi
C 533	36	87.8	145832	8	AC090736	AC090736 Homo sapi	606	36	87.8	167574	14	AC011274	AC011274 Homo sapi
C 534	36	87.8	146177	14	AC155243	AC155243 Mus muscu	C 607	36	87.8	167647	14	AC148024	AC148024 Homo sapi
C 535	36	87.8	146403	14	AC027067	AC027067 Homo sapi	C 608	36	87.8	167792	14	AC113226	AC113226 Papio anu
C 536	36	87.8	147122	8	AC090109	AC090109 Homo sapi	C 609	36	87.8	168818	14	AC148184	AC148184 Macaca mu
C 537	36	87.8	147631	9	AC026617	AC026617 Mus muscu	C 610	36	87.8	168889	9	AL731779	AL731779 Mouse DNA
C 538	36	87.8	147670	8	AC026617	AC026617 Homo sapi	611	36	87.8	169131	8	AC009446	AC009446 Homo sapi
C 539	36	87.8	148027	14	AC132893	AC132893 Mus muscu	C 612	36	87.8	169308	5	BX537251	BX537251 Zebrafish
C 540	36	87.8	150063	14	AC068264	AC068264 Homo sapi	C 613	36	87.8	169890	14	AL662872	AL662872 Homo sapi
C 541	36	87.8	150107	14	AC027582	AC027582 Homo sapi	614	36	87.8	169964	8	AC055710	AC055710 Homo sapi
C 542	36	87.8	150706	14	AC160215	AC160215 Atelelex	615	36	87.8	169969	14	AL390727	AL390727 Homo sapi
C 543	36	87.8	150894	9	AC154190	AC154190 Mus muscu	C 616	36	87.8	170011	8	CNS06C7N	AL390334 Human chr
C 544	36	87.8	151321	14	AC136264	AC136264 Rattus no	617	36	87.8	170321	8	AC097466	AC097466 Homo sapi
C 545	36	87.8	151418	14	AC025100	AC025100 Homo sapi	C 618	36	87.8	170895	14	AC068545	AC068545 Homo sapi
C 546	36	87.8	151475	9	AC153995	AC153995 Mus muscu	C 619	36	87.8	170872	14	AC166141	AC166141 Xenopus t
C 547	36	87.8	151606	15	AP005719	AP005719 Oryza sat	C 620	36	87.8	171099	8	AC092644	AC092644 Homo sapi
C 548	36	87.8	152000	8	AC119749	AC119749 Homo sapi	C 621	36	87.8	171176	8	AC119751	AC119751 Homo sapi
C 549	36	87.8	152010	5	BX511002	BX511002 Zebrafish	C 622	36	87.8	171188	9	AC134431	AC134431 Mus muscu
C 550	36	87.8	152121	14	AP000830	AP000830 Homo sapi	C 623	36	87.8	171640	14	AC152362	AC152362 Atelelex
C 551	36	87.8	152296	8	CR381535	CR381535 Human DNA	624	36	87.8	172077	9	AC149220	AC149220 Mus muscu
C 552	36	87.8	153370	14	AC157852	AC157852 Atelelex	C 625	36	87.8	172239	14	AC079307	AC079307 Homo sapi
C 553	36	87.8	153788	14	AC027472	AC027472 Homo sapi	C 626	36	87.8	172519	14	AC161794	AC161794 Mus muscu
C 554	36	87.8	153860	9	AC101983	AC101983 Mus muscu	627	36	87.8	172535	14	CR735106	CR735106 Danio rer
C 555	36	87.8	155122	14	AC141847	AC141847 Pan trogl	C 628	36	87.8	172649	8	AC079375	AC079375 Homo sapi
C 556	36	87.8	155397	8	AL590523	AL590523 Human DNA	C 629	36	87.8	172832	15	AP004047	AP004047 Oryza sat
C 557	36	87.8	155813	8	AC134878	AC134878 Homo sapi	630	36	87.8	174044	8	AC090675	AC090675 Homo sapi
C 558	36	87.8	155958	14	AC024975	AC024975 Homo sapi	C 631	36	87.8	174316	5	BX572630	BX572630 Zebrafish
C 559	36	87.8	156328	14	AC148280	AC148280 Sorax ara	C 632	36	87.8	175048	14	AC162114	AC162114 Bos tauru
C 560	36	87.8	156825	8	AC026726	AC026726 Homo sapi	C 633	36	87.8	175202	14	AC117872	AC117872 Rattus no
C 561	36	87.8	157091	14	AC025484	AC025484 Homo sapi	C 634	36	87.8	175425	14	AC012208	AC012208 Homo sapi
C 562	36	87.8	157210	5	BX328756	BX328756 Zebrafish	C 635	36	87.8	175742	8	AC021001	AC021001 Homo sapi
C 563	36	87.8	157412	5	CR513790	CR513790 Zebrafish	C 636	36	87.8	175827	8	AL358472	AL358472 Human DNA
C 564	36	87.8	157533	8	AP003383	AP003383 Homo sapi	C 637	36	87.8	176022	14	AC159164	AC159164 Atelelex
C 565	36	87.8	157733	8	AC138771	AC138771 Homo sapi	C 638	36	87.8	177101	14	AC146642	AC146642 Otlemur
C 566	36	87.8	157917	14	AC027224	AC027224 Homo sapi	639	36	87.8	177219	8	HS179D22	HS179D22 Zebrafish
C 567	36	87.8	158069	8	CR392039	CR392039 Human DNA	C 640	36	87.8	177948	9	AC154335	AC154335 Mus muscu
C 568	36	87.8	158738	14	AC015500	AC015500 Homo sapi	C 641	36	87.8	177966	5	CR788312	CR788312 Zebrafish
C 569	36	87.8	158892	14	CNS01R14	AL162871 Homo sapi	C 642	36	87.8	178032	14	AC087869	AC087869 Mus muscu
C 570	36	87.8	158983	14	AC020598	AC020598 Homo sapi	C 643	36	87.8	178451	8	AL139396	AL139396 Human DNA
C 571	36	87.8	159345	14	AC148492	AC148492 Otlemur	644	36	87.8	178657	14	AC116088	AC116088 Rattus no
C 572	36	87.8	159634	14	AC010193	AC010193 Homo sapi	645	36	87.8	179206	8	CNS01DS6	AL121656 BAC seque
C 573	36	87.8	159739	4	AC150679	AC150679 Monodelph	C 646	36	87.8	179508	8	AC010799	AC010799 Homo sapi
C 574	36	87.8	160109	8	AL161727	AL161727 Human DNA	647	36	87.8	179693	8	AL591856	AL591856 Human DNA
C 575	36	87.8	160460	14	AC024374	AC024374 Homo sapi	648	36	87.8	180315	14	AC161187	AC161187 Mus muscu
C 576	36	87.8	161079	14	AC036148	AC036148 Homo sapi	649	36	87.8	180708	8	AL590725	AL590725 Human DNA
C 577	36	87.8	161147	8	AL354822	AL354822 Human DNA	C 650	36	87.8	181249	14	AC148479	AC148479 Zea mays
C 578	36	87.8	161781	14	CR376800	CR376800 Danio rer	651	36	87.8	181419	8	AL445984	AL445984 Human DNA
C 579	36	87.8	162153	8	AC092331	AC092331 Homo sapi	C 652	36	87.8	181510	14	AC020737	AC020737 Homo sapi
C 580	36	87.8	163184	5	BX537310	BX537310 Zebrafish	653	36	87.8	181996	14	AC108614	AC108614 Rattus no
C 581	36	87.8	164147	15	AP005784	AP005784 Oryza sat	C 654	36	87.8	181997	14	AC165399	AC165399 Oryctoleg
C 582	36	87.8	164654	14	AC147952	AC147952 Papio anu	C 655	36	87.8	182261	14	AC026134	AC026134 Homo sapi
C 583	36	87.8	164711	14	AC139564	AC139564 Homo sapi	C 656	36	87.8	182314	14	AC013670	AC013670 Homo sapi
C 584	36	87.8	164830	4	AC144690	AC144690 Bidelphe	657	36	87.8	182943	9	AC101743	AC101743 Mus muscu
C 585	36	87.8	165223	8	AC104059	AC104059 Homo sapi	658	36	87.8	183774	8	AL596217	AL596217 Human DNA
C 586	36	87.8	165277	5	AL954184	AL954184 Zebrafish	C 659	36	87.8	183858	9	AC102452	AC102452 Mus muscu
C 587	36	87.8	165333	8	BX546492	BX546492 Human DNA	660	36	87.8	183914	14	AC069331	AC069331 Homo sapi

661	36	87.8	184424	2	AC010051	AC010051 Drosophil	c 734	36	87.8	202391	9	AC110040	AC110040 Mus muscu
662	36	87.8	184760	9	AL102499	AL102499 Mus muscu	735	36	87.8	202469	5	CR786576	CR786576 Zebrafish
663	36	87.8	184887	8	AL713998	AL713998 Human DNA	736	36	87.8	202568	9	AC009407	AC009407 Homo sapi
664	36	87.8	185024	14	AC027649	AC027649 Mus muscu	737	36	87.8	203071	9	AL606522	AL606522 Mouse DNA
665	36	87.8	185517	9	AC140231	AC140231 Mus muscu	c 738	36	87.8	203138	5	EX640539	EX640539 Zebrafish
666	36	87.8	185644	8	AC087455	AC087455 Homo sapi	739	36	87.8	203169	8	AL390785	AL390785 Human DNA
667	36	87.8	185838	9	AC133177	AC133177 Mus muscu	740	36	87.8	203287	5	CR376756	CR376756 Zebrafish
668	36	87.8	186299	14	EX936301	EX936301 Danio rer	741	36	87.8	203530	8	AC025097	AC025097 Homo sapi
669	36	87.8	186444	2	AC093439	AC093439 Drosophil	c 742	36	87.8	204239	14	CT010494	CT010494 Mus muscu
670	36	87.8	186588	14	AC141218	AC141218 Rattus no	c 743	36	87.8	204413	9	AL645910	AL645910 Mouse DNA
671	36	87.8	186606	9	AC129308	AC129308 Mus muscu	c 744	36	87.8	204798	14	AC125810	AC125810 Rattus no
672	36	87.8	186622	14	AC023917	AC023917 Homo sapi	745	36	87.8	204908	8	AL158163	AL158163 Human DNA
673	36	87.8	186739	8	EX572084	EX572084 Human DNA	746	36	87.8	205028	9	AL154862	AL154862 Mus muscu
674	36	87.8	187110	9	AC131303	AC131303 Mus muscu	747	36	87.8	205524	9	EX678769	EX678769 Mouse DNA
675	36	87.8	187161	14	AC159759	AC159759 Bos tauru	c 748	36	87.8	205759	5	CR589882	CR589882 Zebrafish
676	36	87.8	187431	9	AC115959	AC115959 Mus muscu	c 749	36	87.8	205807	14	CR388418	CR388418 Danio rer
677	36	87.8	187620	15	AP005527	AP005527 Oryza sat	750	36	87.8	205887	9	AC140269	AC140269 Mus muscu
678	36	87.8	187710	9	AC123807	AC123807 Mus muscu	751	36	87.8	206579	2	AC150259	AC150259 Aedes aeg
679	36	87.8	187857	14	AC044850	AC044850 Homo sapi	752	36	87.8	207501	9	AL590389	AL590389 Mouse DNA
680	36	87.8	187984	14	AP001262	AP001262 Homo sapi	c 753	36	87.8	207901	9	AL590389	AL590389 Mouse DNA
681	36	87.8	188076	14	AC166614	AC166614 Oryctolag	c 754	36	87.8	208200	14	AC157026	AC157026 Bos tauru
682	36	87.8	188306	14	CR762398	CR762398 Danio rer	755	36	87.8	208678	14	AC162532	AC162532 Zebrafish
683	36	87.8	188448	8	AL355432	AL355432 Homo sapi	c 756	36	87.8	209698	5	CR381585	CR381585 Zebrafish
684	36	87.8	188491	8	H556F10	H556F10 Human DNA	c 757	36	87.8	209819	9	AC154839	AC154839 Mus muscu
685	36	87.8	188684	8	AL355312	AL355312 Homo sapi	c 758	36	87.8	210013	9	AC134791	AC134791 Mus muscu
686	36	87.8	188776	5	EX537130	EX537130 Zebrafish	c 759	36	87.8	210548	9	AC137905	AC137905 Mus muscu
687	36	87.8	188887	8	AC116364	AC116364 Homo sapi	c 760	36	87.8	211422	8	AC073493	AC073493 Homo sapi
688	36	87.8	189169	14	AC122690	AC122690 Homo sapi	761	36	87.8	211546	14	AC135554	AC135554 Gallus ga
689	36	87.8	189199	8	AC006566	AC006566 Homo sapi	762	36	87.8	211779	9	AC163012	AC163012 Mus muscu
690	36	87.8	189308	8	AC093416	AC093416 Homo sapi	c 763	36	87.8	212131	14	AC142554	AC142554 Pan trogl
691	36	87.8	189348	8	AC131154	AC131154 Homo sapi	764	36	87.8	212174	14	AC105896	AC105896 Rattus no
692	36	87.8	189648	14	AC069163	AC069163 Homo sapi	765	36	87.8	212549	8	AC009492	AC009492 Homo sapi
693	36	87.8	190286	14	AC147591	AC147591 Papio anu	766	36	87.8	212659	14	AL390737	AL390737 Homo sapi
694	36	87.8	190708	8	AC005261	AC005261 Homo sapi	767	36	87.8	213227	5	BX572077	BX572077 Zebrafish
695	36	87.8	191109	14	AC027456	AC027456 Homo sapi	c 768	36	87.8	213513	5	BX572077	BX572077 Zebrafish
696	36	87.8	191144	8	AC138773	AC138773 Homo sapi	769	36	87.8	213551	9	AC156976	AC156976 Bos tauru
697	36	87.8	191239	14	AP000883	AP000883 Homo sapi	c 770	36	87.8	213575	9	AC154438	AC154438 Mus muscu
698	36	87.8	191713	14	CR925834	CR925834 Danio rer	c 771	36	87.8	213708	14	AC157045	AC157045 Bos tauru
699	36	87.8	191996	6	Q873070	Q873070 Sequence	c 772	36	87.8	214411	14	AC069297	AC069297 Mus muscu
700	36	87.8	192016	14	AC026540	AC026540 Homo sapi	c 773	36	87.8	214523	8	AP005900	AP005900 Homo sapi
701	36	87.8	192068	14	AC036208	AC036208 Homo sapi	c 774	36	87.8	215521	14	AL126103	AL126103 Rattus no
702	36	87.8	192299	9	AC155301	AC155301 Mus muscu	775	36	87.8	215611	14	AL359768	AL359768 Homo sapi
703	36	87.8	192406	9	EX005086	EX005086 Mouse DNA	776	36	87.8	216200	8	BS000123	BS000123 Pan trogl
704	36	87.8	193375	9	AC136020	AC136020 Mus muscu	c 777	36	87.8	216368	14	CT025592	CT025592 Homo sapi
705	36	87.8	193665	8	AC105314	AC105314 Homo sapi	c 778	36	87.8	216571	8	AC009361	AC009361 Homo sapi
706	36	87.8	193681	8	AC069382	AC069382 Homo sapi	779	36	87.8	216751	14	AC137457	AC137457 Rattus no
707	36	87.8	193978	9	AC114536	AC114536 Mus muscu	c 780	36	87.8	218105	14	AC099238	AC099238 Rattus no
708	36	87.8	194074	9	AC150894	AC150894 Mus muscu	c 781	36	87.8	218258	9	AC096310	AC096310 Rattus no
709	36	87.8	194111	15	AP005160	AP005160 Oryza sat	782	36	87.8	218479	14	AC153030	AC153030 Bos tauru
710	36	87.8	194122	14	AC138778	AC138778 Homo sapi	c 783	36	87.8	218684	14	AC159870	AC159870 Bos tauru
711	36	87.8	194221	9	AC139863	AC139863 Mus muscu	784	36	87.8	218908	14	CT010471	CT010471 Mus muscu
712	36	87.8	194310	8	AP001267	AP001267 Homo sapi	785	36	87.8	219940	14	AC087866	AC087866 Mus muscu
713	36	87.8	194449	14	CR677617	CR677617 Danio rer	786	36	87.8	220260	6	AC161138	AC161138 Bos tauru
714	36	87.8	194560	14	AC146623	AC146623 Mus muscu	c 787	36	87.8	220281	8	AC138777	AC138777 Homo sapi
715	36	87.8	194636	14	CT010577	CT010577 Mus muscu	c 788	36	87.8	220371	14	AC109966	AC109966 Rattus no
716	36	87.8	194744	14	CR936366	CR936366 Danio rer	c 789	36	87.8	220892	14	EX936393	EX936393 Danio rer
717	36	87.8	194937	8	BS000153	BS000153 Pan trogl	790	36	87.8	221264	9	AC155258	AC155258 Mus muscu
718	36	87.8	195087	5	EX000438	EX000438 Zebrafish	791	36	87.8	221568	14	AC103452	AC103452 Rattus no
719	36	87.8	195141	8	CR80508P	CR80508P Human chr	c 792	36	87.8	221612	9	AC164250	AC164250 Mus muscu
720	36	87.8	197245	14	AL591164	AL591164 Homo sapi	793	36	87.8	221981	14	AC103165	AC103165 Rattus no
721	36	87.8	197770	14	EX322632	EX322632 Homo sapi	c 794	36	87.8	222787	2	AE003637	AE003637 Drosophil
722	36	87.8	197946	9	AC110251	AC110251 Mus muscu	c 795	36	87.8	223192	14	AC127011	AC127011 Rattus no
723	36	87.8	198312	14	AC141368	AC141368 Rattus no	c 796	36	87.8	223425	5	EX247888	EX247888 Zebrafish
724	36	87.8	198343	14	AC165288	AC165288 Mus muscu	797	36	87.8	224203	14	AC109957	AC109957 Rattus no
725	36	87.8	198900	9	AC123556	AC123556 Mus muscu	c 798	36	87.8	224429	14	AC12452	AC12452 Rattus no
726	36	87.8	199051	14	AC156258	AC156258 Callithri	799	36	87.8	225962	14	AC094695	AC094695 Rattus no
727	36	87.8	199198	8	AC021413	AC021413 Homo sapi	800	36	87.8	226259	14	AC097623	AC097623 Rattus no
728	36	87.8	199370	14	AC141908	AC141908 Rattus no	801	36	87.8	226438	14	AC132793	AC132793 Rattus no
729	36	87.8	200476	8	AC079271	AC079271 Homo sapi	c 802	36	87.8	226657	9	AL731843	AL731843 Mouse DNA
730	36	87.8	201124	14	AC074287	AC074287 Homo sapi	c 803	36	87.8	226781	9	AL672230	AL672230 Mouse DNA
731	36	87.8	201246	9	AL683894	AL683894 Mouse DNA	c 804	36	87.8	226924	9	AC113490	AC113490 Mus muscu
732	36	87.8	201709	8	EX004987	EX004987 Human DNA	805	36	87.8	227763	14	AC126518	AC126518 Rattus no
733	36	87.8	202046	9	AL837516	AL837516 Mouse DNA	806	36	87.8	227844	14	AC138049	AC138049 Canis fam

C 807	36	87.8	227922	9	AC120547	AC120547 Mus muscu	C 880	36	87.8	292173	14	BX294381	BX294381 Danio rer
C 808	36	87.8	227932	14	AC158060	AC158060 Bos tauru	C 881	36	87.8	300029	1	AE015936	AE015936 Clostridi
C 809	36	87.8	228013	14	AC105582	AC105582 Rattus no	C 882	36	87.8	301439	1	AE015943	AE015943 Clostridi
C 810	36	87.8	228052	14	AC128624	AC128624 Rattus no	C 883	36	87.8	302341	1	AE017163	AE017163 Prochloro
811	36	87.8	228932	9	AL645669	AL645669 Mouse DNA	884	36	87.8	304384	14	AC120944	AC120944 Rattus no
C 812	36	87.8	228933	9	AC011661	AC011661 Mus muscu	885	36	87.8	305833	8	AY275833	AY275833 Homo sapi
C 813	36	87.8	229012	14	AC021084	AC021084 Homo sapi	886	36	87.8	305724	14	AC128763	AC128763 Rattus no
C 814	36	87.8	229118	14	AC103306	AC103306 Rattus no	887	36	87.8	310520	14	AC152330	AC152330 Bos tauru
C 815	36	87.8	229171	14	AC126947	AC126947 Rattus no	C 888	36	87.8	313157	2	AE003593	AE003593 Drosophil
C 816	36	87.8	229179	14	AC130231	AC130231 Rattus no	C 889	36	87.8	313264	14	AC023053	AC023053 Homo sapi
817	36	87.8	231014	5	BX784031	BX784031 Zebrafish	C 890	36	87.8	317022	1	AE017148	AE017148 Helicobac
C 818	36	87.8	231406	14	AC132626	AC132626 Rattus no	C 891	36	87.8	320135	14	AC151850	AC151850 Takifugu
C 819	36	87.8	231853	14	AC093352	AC093352 Mus muscu	C 892	36	87.8	327261	14	AC105726	AC105726 Rattus no
C 820	36	87.8	232058	14	AC094131	AC094131 Rattus no	C 893	36	87.8	330050	2	PFA929355	AL929355 Plasmodiu
C 821	36	87.8	232133	14	AC095770	AC095770 Rattus no	C 894	36	87.8	331666	14	AC133245	AC133245 Rattus no
C 822	36	87.8	232392	13	AF250284	AF250284 Anisacta m	C 895	36	87.8	332889	14	AC098450	AC098450 Rattus no
C 823	36	87.8	232814	14	AC159521	AC159521 Bos tauru	C 896	36	87.8	340000	8	AP001729	AP001729 Homo sapi
C 824	36	87.8	232923	14	AC094246	AC094246 Rattus no	C 897	36	87.8	347582	2	PFWAL4P1	AL034557 Plasmodiu
825	36	87.8	232943	14	AC121451	AC121451 Rattus no	C 898	36	87.8	349533	14	AC132500	AC132500 Rattus no
C 826	36	87.8	233010	14	AC096368	AC096368 Rattus no	C 899	36	87.8	349742	1	BX572090	BX572090 Prochloro
C 827	36	87.8	233278	14	AC134499	AC134499 Rattus no	C 900	35	85.4	173	10	BV101072	BV101072 RFAMSEQO
828	36	87.8	233326	9	AC111111	AC111111 Mus muscu	C 901	35	85.4	197	13	AF482001	AF482001 Canine co
C 829	36	87.8	233366	14	AC094867	AC094867 Rattus no	C 902	35	85.4	242	5	AY127562	AY127562 Callipepl
C 830	36	87.8	234843	14	AC106580	AC106580 Rattus no	C 903	35	85.4	266	6	AX341129	AX341129 Sequence
C 831	36	87.8	234922	14	AC097055	AC097055 Rattus no	C 904	35	85.4	266	6	AX341306	AX341306 Sequence
C 832	36	87.8	235664	14	AC125522	AC125522 Rattus no	C 905	35	85.4	288	9	MMPOLE18	AF126394 Mus muscu
C 833	36	87.8	237320	14	AC141339	AC141339 Rattus no	C 906	35	85.4	345	13	AY243147	AY243147 HIV-1 iso
C 834	36	87.8	238387	14	AC103224	AC103224 Rattus no	C 907	35	85.4	354	10	CR379831	CR379831 Arabidops
C 835	36	87.8	239047	14	AC163025	AC163025 Mus muscu	C 908	35	85.4	356	15	AY702436	AY702436 Solenopho
C 836	36	87.8	239507	14	AC136588	AC136588 Rattus no	C 909	35	85.4	371	15	AB164373	AB164373 Trichospo
C 837	36	87.8	240000	14	AC158688	AC158688 Bos tauru	C 910	35	85.4	391	15	AY364294	AY364294 Solenopho
C 838	36	87.8	240742	14	AC096830	AC096830 Rattus no	C 911	35	85.4	408	6	CQ723592	CQ723592 Sequence
C 839	36	87.8	240914	14	AC097252	AC097252 Rattus no	C 912	35	85.4	412	6	AX778488	AX778488 Sequence
C 840	36	87.8	241394	14	AC132789	AC132789 Rattus no	C 913	35	85.4	427	10	BV007002	BV007002 8A09F09.x
841	36	87.8	242467	14	AC128990	AC128990 Rattus no	C 914	35	85.4	459	6	AR272335	AR272335 Sequence
C 842	36	87.8	242701	14	AC122596	AC122596 Rattus no	C 915	35	85.4	459	6	AR275916	AR275916 Sequence
843	36	87.8	243118	14	AC157397	AC157397 Bos tauru	C 916	35	85.4	459	6	AR406191	AR406191 Sequence
C 844	36	87.8	243587	14	AL591179	AL591179 Homo sapi	C 917	35	85.4	459	6	AR440041	AR440041 Sequence
C 845	36	87.8	244524	9	AL543380	AL543380 Mus muscu	C 918	35	85.4	459	6	AR472200	AR472200 Sequence
C 846	36	87.8	245531	6	CQ870385	CQ870385 Sequence	C 919	35	85.4	459	6	AR542852	AR542852 Sequence
847	36	87.8	245855	14	AC106203	AC106203 Rattus no	C 920	35	85.4	459	6	AR638010	AR638010 Sequence
C 848	36	87.8	246210	14	AC128400	AC128400 Rattus no	C 921	35	85.4	459	6	AX062420	AX062420 Sequence
C 849	36	87.8	246259	14	AC128274	AC128274 Rattus no	C 922	35	85.4	459	6	AX367337	AX367337 Sequence
C 850	36	87.8	246401	14	AC094195	AC094195 Rattus no	C 923	35	85.4	527	6	AR226592	AR226592 Sequence
C 851	36	87.8	247381	14	AC134193	AC134193 Rattus no	C 924	35	85.4	546	10	G73505	G73505 YAC185right
C 852	36	87.8	248002	14	AC159043	AC159043 Bos tauru	C 925	35	85.4	549	10	BV418888	BV418888 S229P6490
C 853	36	87.8	249307	14	AC095208	AC095208 Rattus no	C 926	35	85.4	571	10	BV038849	BV038849 S212P6856
C 854	36	87.8	249574	14	AC094707	AC094707 Rattus no	C 927	35	85.4	580	10	BV247045	BV247045 S234P6156
C 855	36	87.8	249850	14	AC150413	AC150413 Branchios	C 928	35	85.4	615	6	CQ736619	CQ736619 Sequence
C 856	36	87.8	251251	14	AC130500	AC130500 Rattus no	C 929	35	85.4	617	10	BV360414	BV360414 S231P6204
C 857	36	87.8	252248	14	AC094553	AC094553 Rattus no	C 930	35	85.4	619	10	BV307461	BV307461 S236P6373
C 858	36	87.8	252813	14	AC118528	AC118528 Rattus no	C 931	35	85.4	624	10	BV339296	BV339296 S230P6206
C 859	36	87.8	252989	14	AC126984	AC126984 Rattus no	C 932	35	85.4	627	10	BV390745	BV390745 S243P685F
C 860	36	87.8	253033	14	AC112741	AC112741 Rattus no	C 933	35	85.4	632	10	BV057314	BV057314 S212P6924
C 861	36	87.8	253342	14	AC097968	AC097968 Rattus no	C 934	35	85.4	642	6	AX954129	AX954129 Sequence
C 862	36	87.8	253453	14	AC111906	AC111906 Rattus no	C 935	35	85.4	653	10	BV354348	BV354348 S230P63FB
C 863	36	87.8	253770	14	AC156810	AC156810 Bos tauru	C 936	35	85.4	665	10	BV018302	BV018302 S212P6725
C 864	36	87.8	255569	14	AC112354	AC112354 Rattus no	C 937	35	85.4	701	6	AR206134	AR206134 Sequence
C 865	36	87.8	256473	14	AC109689	AC109689 Rattus no	C 938	35	85.4	701	6	AR372882	AR372882 Sequence
C 866	36	87.8	257451	14	AC098508	AC098508 Rattus no	C 939	35	85.4	716	10	BV069529	BV069529 S212P6063
C 867	36	87.8	257660	14	AC121483	AC121483 Rattus no	C 940	35	85.4	720	6	CQ645856	CQ645856 Sequence
C 868	36	87.8	259772	14	AC106943	AC106943 Rattus no	C 941	35	85.4	729	6	AX381944	AX381944 Sequence
C 869	36	87.8	259795	14	AC128267	AC128267 Rattus no	C 942	35	85.4	765	6	AR206126	AR206126 Sequence
C 870	36	87.8	261124	14	AC098233	AC098233 Rattus no	C 943	35	85.4	765	6	AR372874	AR372874 Sequence
C 871	36	87.8	263876	8	AC019233	AC019233 Homo sapi	C 944	35	85.4	768	1	VPO580876	AT580876 Vibrio po
C 872	36	87.8	265750	14	AC110831	AC110831 Rattus no	C 945	35	85.4	768	1	AJ842497	AJ842497 Vibrio po
C 873	36	87.8	267585	14	CR352249	CR352249 Danio rer	C 946	35	85.4	773	10	BV615026	BV615026 S217P6355
C 874	36	87.8	267710	14	AC161793	AC161793 Mus muscu	C 947	35	85.4	790	10	BV603047	BV603047 S216P6021
C 875	36	87.8	272065	14	AC117029	AC117029 Rattus no	C 948	35	85.4	790	10	BV493704	BV493704 S217P6673
C 876	36	87.8	275606	14	AC123139	AC123139 Rattus no	C 949	35	85.4	797	1	AJ842498	AJ842498 Vibrio po
C 877	36	87.8	276814	14	AC097924	AC097924 Rattus no	C 950	35	85.4	825	11	AY890789	AY890789 Synthetic
C 878	36	87.8	288360	14	AC157114	AC157114 Bos tauru	C 951	35	85.4	825	11	BT019955	BT019955 Synthetic
C 879	36	87.8	289087	14	AC110678	AC110678 Rattus no	C 952	35	85.4	837	10	BV648501	BV648501 S215P6028

Alignment Scores:
Pred. No.: 345 Length: 901
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-774-176-8 (1-9) x AX025013 (1-901)

QY 1 AlaIlePheLeuLeuValLeuTyrLeu 9
DB 669 GCCATCTTCCTACTGGTTTGATTGG 695

RESULT 5
AX316088
LOCUS AX316088 901 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 3 from Patent EP1160323.
ACCESSION AX316088
VERSION AX316088.1 GI:17899280
KEYWORDS
SOURCE Canis sp.
ORGANISM Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE
1
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE St4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 3 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
LOCATION/Qualifiers
FEATURES
source 1..901
/mol_type="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN

Alignment Scores:
Pred. No.: 345 Length: 901
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-774-176-8 (1-9) x AX316088 (1-901)

QY 1 AlaIlePheLeuLeuValLeuTyrLeu 9
DB 669 GCCATCTTCCTACTGGTTTGATTGG 695

RESULT 6
AX829164
LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1
AUTHORS Mack, D.H., Gish, K.C. and Afari, D.
TITLE Methods of diagnosis of breast cancer, compositions and methods of
screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)
LOCATION/Qualifiers
FEATURES
source 1..927
/mol_type="Homo sapiens"

```
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 354 Length: 927
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x AX829164 (1-927)

Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 760 GCTATTTTCCTCGTGGTTTGATTTG 786

RESULT 7
AX467373 1260 bp DNA linear PAT 16-JUL-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
Felis sp.
Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/mol_type="unassigned DNA"
/db_xref="taxon:9687"

ORIGIN
Alignment Scores:
Pred. No.: 456 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x AX467373 (1-1260)

Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 1099 GCCATTTTCTACTGGTTTGACTTG 1125

RESULT 8
AX821533 1260 bp DNA linear PAT 10-DEC-2003
LOCUS
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
TITLE MHC class I peptide epitopes from the human 5t4 tumor-associated
antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;

/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN
Alignment Scores:
Pred. No.: 456 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x AX821548 (1-1260)

Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 1099 GCCATTTTCTACTGGTTTGACTTG 1125

RESULT 10
BD249731 1263 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 (bases 1 to 1263)
```

AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL,KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
PC C07K19/00,
PC C12N15/00,
CC Polypeptide
FT Key Location/Qualifiers
FT source 1..1263
FT /organism='Homo sapiens (human)'.
FEATURES source
1..1263 Location/Qualifiers
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores: 457 Length: 1263
Pred. No.: 41.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-8 (1-9) x BD249731 (1-1263)
Qy 1 Alal1ePheLeuLeuValLeuTyrLeu 9
Db 1102 GCTATTTCCTCTGCTTTGTATTG 1128
RESULT 11
AX025011
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES source
1..1263 Location/Qualifiers
/organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores: 457 Length: 1263
Pred. No.: 41.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-8 (1-9) x AX025011 (1-1263)

Qy 1 Alal1ePheLeuLeuValLeuTyrLeu 9
Db 1102 GCTATTTCCTCTGCTTTGTATTG 1128
RESULT 12
AX149553
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
Billard,P.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES source
1..1263 Location/Qualifiers
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/note='5T4'
ORIGIN
Alignment Scores: 457 Length: 1263
Pred. No.: 41.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-8 (1-9) x AX149553 (1-1263)
Qy 1 Alal1ePheLeuLeuValLeuTyrLeu 9
Db 1102 GCACCTTCCTACTGCTTTGTATTG 1128
RESULT 13
AX316086
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES source
1..1263 Location/Qualifiers
/organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores: 457 Length: 1263
Pred. No.: 41.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0%

```
DB: 6 Gaps: 0
US-10-774-176-8 (1-9) x AX316086 (1-1263)
Qy 1 Alal1ePheLeuVal1euTyr1eu 9
Db 1102 GCTATTTCCTCCTCGTTTGTATTG 1128
RESULT 14
LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO238612.
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
Location/Qualifiers
1..1263
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"
ORIGIN
Alignment Scores:
Pred. No.: 457 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-774-176-8 (1-9) x AX467371 (1-1263)
Qy 1 Alal1ePheLeuVal1euTyr1eu 9
Db 1102 GCCATCTTCTACTCGTTTGTATTG 1128
RESULT 15
LOCUS BD249732 1281 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249732
VERSION BD249732.1 GI:33059502
KEYWORDS JP 2002530060-A/2.
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 2 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Mus musculus (mouse)
PN JP 2002530060-A/2
PD 17-SEP-2002
PP 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09, A61K39/00, A61K48/00, C07K7/06, C07K14/065,
PC C07K19/00,
PC C12N15/00
```

```
CC Polypeptide Location/Qualifiers
FH Key 1..1281
FT source /organism="Mus musculus (mouse)".
FEATURES
source
Location/Qualifiers
1..1281
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.: 462 Length: 1281
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-774-176-8 (1-9) x BD249732 (1-1281)
Qy 1 Alal1ePheLeuVal1euTyr1eu 9
Db 1120 GCTATTTCCTCCTCGTTTGTATTG 1146
RESULT 16
LOCUS AX025012 1281 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 2 from Patent WO029428.
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 2 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source
Location/Qualifiers
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.: 462 Length: 1281
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-774-176-8 (1-9) x AX025012 (1-1281)
Qy 1 Alal1ePheLeuVal1euTyr1eu 9
Db 1120 GCTATTTCCTCCTCGTTTGTATTG 1146
RESULT 17
LOCUS AX316087 1281 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 2 from Patent EP160323.
ACCESSION AX316087
VERSION AX316087.1 GI:17899279
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

1
REFERENCE
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 2 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 462 Length: 1281
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x AX316087 (1-1281)

Qy 1 AlallePheLeuLeuValleuTyrieu 9
Db 1120 GCTATTTCCTCCTCGTTTGATTG 1146

RESULT 18
LOCUS CQ731678 2053 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 17612 from Patent WO02068579.
ACCESSION CQ731678
VERSION CQ731678.1 GI:42308932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, B.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 17612 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..2053
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 682 Length: 2053
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x CQ731678 (1-2053)

Qy 1 AlallePheLeuLeuValleuTyrieu 9
Db 1188 GCTATTTCCTCCTCGTTTGATTG 1214

RESULT 19
LOCUS HST40A 2053 bp RNA linear PRI 18-APR-2005
DEFINITION Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.

229083
ACCESSION 229083.1 GI:435654
KEYWORDS 5T4 gene; 5T4 oncofoetal antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 2053)
REFERENCE
AUTHORS Myers, K.A., Rahi-Saund, V., Davison, M.D., Young, J.A., Cheater, A.J.
and Stern, P.L.
TITLE Isolation of a cDNA encoding 5T4 oncofoetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
JOURNAL J. Biol. Chem. 269 (12), 9319-9324 (1994)
PUBMED 813670
REFERENCE 2 (bases 1 to 2053)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
FEATURES
source
1..2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label=N-flank
85..1347
/codon_start=1
/evidence=experimental
/product="5T4 oncofoetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGRGPAGDGRRLRLARLALVLGWSSSPPTSSASFSS
APFLASAVSAQPLPDQCPCSEARATKVCNRLNTEVPTDLPAYVRLPLTGNQ
LAVLPAGAPARPPPLAELALNLSSGRDLVRAGAPEHLPSLRQLDLSHNLADLSPP
AFSGSNASVAPSPLVELLNHIVPPEDERQNRSEFGMVVAALLAGRALQGLRLELA
SNHFLYLPDVLQPLSLRDLNNSLSVLTYSFRNLTHLSLHEDNALKVRLNG
TLAEIQLGLPHIRVFLDNNPWCDCNMDMTWLKETEYVQGGKDRLLTCAYPEKMRNRYL
LELNSADLDCDPLPPSLQTSYVFLGIVLALIGAIFLLVLYLNKRGTKKWMHNRDACC
RDHMEGYHYRYEINADPRLTNLSSNDV"
130..171
sig_peptide
misc_RNA 373..966
/product="Leucine rich repeat region"
/label=LRRS
966..1119
misc_RNA /product="LRR C-terminal flank"
/label=C-flank
1153..1215
misc_RNA /product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

ORIGIN

Alignment Scores:
Pred. No.: 682 Length: 2053
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-8 (1-9) x H9ST40A (1-2053)

Qy 1 AlaiIlePheLeuValLeuTyrLeu 9
Db 1186 GCTATTTCCTCCTCGTTTGATTTG 1212

RESULT 20
AF063939
LOCUS
DEFINITION Rattus norvegicus 574 oncofetal antigen homolog (574) mRNA, ROD 01-JAN-2000
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 2333)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
TITLE Ninkina, N. and Buchman, V.L.
AUTHORS Structure and expression of the rat 574 gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2333)
AUTHORS Buchman, V.L.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1998) School of Biomedical Sciences, University of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS, UK

FEATURES
source
1..2333
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue type="cerebellum"
/dev stage="newborn"

gene
1..2333
/gene="574"

5'UTR
1..363
/gene="574"

CDS
364..1644
/gene="574"

/codon_start=1
/product="574 oncofetal antigen homolog"

/protein_id="AAF21770.1"

/db_xref="GI:6650212"

/translation="MPGAGSRGSPAGDGLRLRLALVLLGWVSAPSSLPSSSTS
PAAFLASGAPPPAPRCPAACCESEAAATVKVNRNLLEVPADLPVYVRLPLTNGQ
MTVLPAQAFAPQPLADLVNLGNHLKEVGAGAFHLPGRLRLDLSHNPILNSAF
TFAGSVSVTFSELELLILHIVPPEDQRQNGSEGVAFEGVAAALRGLALRGL
HLELASNHLYLPRLDLQPLSKHLDRNNLSVLTYSFRNLTHLSLHEDNAL
KVLHNSTLARGQLARVFLDNNPWCDVMWMLKETEVPDKARLTCAPEK
MNRGLLDTSLDLCDDATLPQSLQTSYVFLGIVLALIGATPFLVLRNKGKKKWH
NIRACRDHMGSHYRYEINADPSLTNLSSSDV"

3'UTR
1645..2333
/gene="574"

polyA_signal
2315..2320
/gene="574"

ORIGIN

Alignment Scores:
Pred. No.: 758 Length: 2333
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-8 (1-9) x AF063939 (1-2333)

Qy 1 AlaiIlePheLeuValLeuTyrLeu 9
Db 1483 GCTATTTCCTCCTCGTTTGATTTG 1509

RESULT 21
BD127282

LOCUS
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127282
VERSION BD127282.1 GI:23222227
KEYWORDS JP 2002017375-A/2713.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2359)

AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL HELIX RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)

PN JP 2002017375-A/2713

PD 22-JAN-2002

PF 07-JUL-2000 JP 2000253172

PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO

PI ISHII,

PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI

SHINICHI KOJIMA,

PI TETSUJI OTSUKI, HISASHI KOGA

PC

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC

10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

PC Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers (424)..(1572).

FEATURES
source
1..2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 765 Length: 2359
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x BD127282 (1-2359)

Qy 1 AlaiIlePheLeuValLeuTyrLeu 9

Db 1525 GCTATTTCCTCCTCGTTTGATTTG 1551

RESULT 22

LOCUS
DEFINITION Sequence 2864 from Patent EP1396543.
ACCESSION CQ782724
VERSION CQ782724.1 GI:45502667

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.


```

TITLE
JOURNAL
Primer for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 2864 10-MAR-2004; (JP)
Research Association for Biotechnology (JP)
FEATURES
source
1. .2359
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"
/translatability="WPGSGRGPAGDGRRLRLARLALVLLGWSSSPSSPSSSPSS
APFLASVSAQPLPDPCCALCESEARVCKVNRNLTEVPDLDLPAYVRNLFLLGNQ
LAVLPAVAFAPAPLAPLALNLSGRDLDEVRAGAFHLPFLQDLNHLNPLADSPF
AFSGNASVAPSPVAVELNHLNVPEDERQNRPFEGMVVAALLAGRLQGLRLLELA
SNHFLYLPDVLQALPDLNHLNLSVSLVTVSPNLTLSLHLELDNALKLVHNG
TLAEIQLGLPFIHVPFLDNNPVCDCMADMTWLKETEVEVQKDRLTCAYPEKMRNVL
LELSADLDCDPLPSPSLQTSYFVLGIVLALGAIFLLVLYLNKGIKK"
ORIGIN
Alignment Scores:
Pred. No.: 765 Length: 2359
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-8 (1-9) x CQ782724 (1-2359)
QY 1 AlaiilePheLeuValLeuTyrLeu 9
|||||
Db 1525 GCTATTTCTCTCGTTTGATTG 1551
RESULT 23
AK074786
LOCUS
2359 bp mRNA linear PRI 03-SBP-2002
DEFINITION
Homo sapiens cDNA FLJ90305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 574 oncofetal trophoblast glycoprotein gene.
ACCESSION
AK074786
VERSION
AK074786.1 GI:22760460
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1
Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahara,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2359)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1. .2359
source
TITLE
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2359)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1. .2359
source
TITLE
JOURNAL
Primer for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 2864 10-MAR-2004; (JP)
Research Association for Biotechnology (JP)
FEATURES
source
1. .2359
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"
/translatability="WPGSGRGPAGDGRRLRLARLALVLLGWSSSPSSPSSSPSS
APFLASVSAQPLPDPCCALCESEARVCKVNRNLTEVPDLDLPAYVRNLFLLGNQ
LAVLPAVAFAPAPLAPLALNLSGRDLDEVRAGAFHLPFLQDLNHLNPLADSPF
AFSGNASVAPSPVAVELNHLNVPEDERQNRPFEGMVVAALLAGRLQGLRLLELA
SNHFLYLPDVLQALPDLNHLNLSVSLVTVSPNLTLSLHLELDNALKLVHNG
TLAEIQLGLPFIHVPFLDNNPVCDCMADMTWLKETEVEVQKDRLTCAYPEKMRNVL
LELSADLDCDPLPSPSLQTSYFVLGIVLALGAIFLLVLYLNKGIKK"
ORIGIN
Alignment Scores:
Pred. No.: 765 Length: 2359
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-8 (1-9) x AK074786 (1-2359)
QY 1 AlaiilePheLeuValLeuTyrLeu 9
|||||
Db 1525 GCTATTTCTCTCGTTTGATTG 1551
RESULT 24
BD127283
LOCUS
2361 bp DNA linear PAT 18-SEP-2002
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD127283
VERSION
BD127283.1 GI:23222228
KEYWORDS
JP 2002017375-A/2714.
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 2361)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 2714 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2714
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUO OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,
C12N5/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N5/00 CC
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers
(426)..(1685).
FT CDS
Location/Qualifiers
1. .2361
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 765 Length: 2361
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0

```

```

Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x BD127283 (1-2361)

QY 1 AlaiilePheLeuValLeuTyrLeu 9
DB 1527 GCTATTTCCTCTGCTTTGTTGTTG 1553

RESULT 25
CQ782726 2361 bp DNA linear PAT 17-MAR-2004
LOCUS
DEFINITION Sequence 2866 from Patent EP1396543.
ACCESSION CQ782726
VERSION CQ782726.1 GI:45502669
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2866 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source Location/Qualifiers
1..2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
426..1688
/notes="unnamed protein product"
/codon_start=1
/db_xref="GI:40881327"

CDS
Pred. No.: 765 Length: 2361
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x AX961916 (1-2361)

QY 1 AlaiilePheLeuValLeuTyrLeu 9
DB 1527 GCTATTTCCTCTGCTTTGTTGTTG 1553

RESULT 27
CQ782726 2361 bp mRNA linear PRI 09-JUL-2005
LOCUS
DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens ST4 oncofetal trophoblast glycoprotein gene.
ACCESSION AK074790
VERSION AK074790.1 GI:22760466
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 Otsuki, T., Ota, T., Nishikawa, T., Hayashi, K., Suzuki, Y.,
Yamamoto, J., Wakamatsu, A., Kimura, K., Sakamoto, K., Hatanaka, N.,
Kawai, Y., Ishii, S., Saito, K., Kojima, S., Sugiyama, T., Ono, T.,
Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A.,
Okumura, K., Nagai, K., Sugano, S. and Isogai, T.
TITLE Signal Sequence and Keyword Trap in silico for Selection of
Full-Length Human cDNAs Encoding Secretion or Membrane Proteins
from Oligo-Capped cDNA Libraries
JOURNAL DNA Res. 12, 117-126 (2005)
REFERENCE
2 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahara, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2361)

```

REMARK
COMMENT

CONTENTS

FEATURES

F

Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.O. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2423)
 Strausberg,R.
 Direct Submission
 Submitted (15-SEP-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hri.nih.gov
 Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
 Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,K., Stantripop,S., Thomas,P.J., Touchman,J.W.,
 Tsugeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILML at: <http://image.llnl.gov>
 Series: IRAC Plate: 123 Row: p Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

Source

1..2423
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:68145 IMAGE:5353871"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

gene

1..2423
 /gene="Tpbp"
 /note="synonym: 574"
 /db_xref="GeneID:21983"
 /db_xref="MGI:1341264"
 402..1682
 /gene="Tpbp"
 /codon_start=1
 /product="Tpbp protein"
 /protein_id="AAH58198.1"
 /db_xref="GI:34849574"
 /db_xref="GeneID:21983"
 /db_xref="MGI:1341264"

CDS

1..2423
 /translation="MPGAGSGPSAGDRLRLARLALVLLGWVSAPSSVPSSTS
 PAAFLASGSAQPPPAECSEARTKVCNRLNLLVPADLPYVRLPLTGNQ
 MTVLPAGARQPPPLADLALNSGNHKEVCAGAFELPGLRLDLSHNPPLTNSAF
 APAGSNASVSPSPLEELIINHIVPPEDQONGSPEGWVAFEGWVAALRSGIALRGL
 TCLEASNHFLPLRDLIAQLPSRLYLDLNNLSVLTYSFNLTHLSLHEDNAL
 KVLHNSTLAEWGSLAHVKVFLDNNPWVCDYMDMVAWLKETEVPDKARLTCAFPK
 MNRGLDLSNDDCDVLPQSLOTSTYVFLGIVLALIGAILLVLYLNKRGIKKWMH
 NTRDADRHMEGHYRYEINADPRLTNLSNSDV"

misc_feature

642..1262
 /gene="Tpbp"
 /note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
 protein [Function unknown]"
 /db_xref="CDD:COG4886"
 1299..1415
 /gene="Tpbp"
 /note="LRRCT; Region: Leucine rich repeat C-terminal
 domain"
 /db_xref="CDD:smart00082"

misc_feature

642..1262
 /gene="Tpbp"
 /note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
 protein [Function unknown]"
 /db_xref="CDD:COG4886"
 1299..1415
 /gene="Tpbp"
 /note="LRRCT; Region: Leucine rich repeat C-terminal
 domain"
 /db_xref="CDD:smart00082"

ORIGIN

Alignment Scores:
 Pred. No.: 782 Length: 2423
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-8 (1-9) x BC058198 (1-2423)

Qy 1 AlailePheLeuLeuValLeuTyrLeu 9

Db 1521 GCTATTTCCTCTCTCGTTTGTATTG 1547

RESULT 31

AX961912
 LOCUS
 DEFINITION Sequence 123 from Patent WO03104277.
 ACCESSION AX961912
 VERSION AX961912.1 GI:40881322
 KEYWORDS
 SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1
 Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
 Stat6 activation gene
 Patent: WO 03104277-A 123 18-DEC-2003;
 Asahi Kasei Kabushiki Kaisha (JP)

FEATURES

Location/Qualifiers
 source
 1..2557
 /organism="Mus musculus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10090"
 556..1836
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="GI:40881323"
 /db_xref="GI:40881323"
 /translation="MPGAGSGPSAGDRLRLARLALVLLGWVSAPSSVPSSTS
 PAAFLASGSAQPPPAECSEARTKVCNRLNLLVPADLPYVRLPLTGNQ
 MTVLPAGARQPPPLADLALNSGNHKEVCAGAFELPGLRLDLSHNPPLTNSAF
 VPAGSNASVSPSPLEELIINHIVPPEDQONGSPEGWVAFEGWVAALRSGIALRGL
 TCLEASNHFLPLRDLIAQLPSRLYLDLNNLSVLTYSFNLTHLSLHEDNAL
 KVLHNSTLAEWGSLAHVKVFLDNNPWVCDYMDMVAWLKETEVPDKARLTCAFPK
 MNRGLDLSNDDCDVLPQSLOTSTYVFLGIVLALIGAILLVLYLNKRGIKKWMH
 NTRDADRHMEGHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
 Pred. No.: 817 Length: 2557
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-8 (1-9) x AX961912 (1-2557)

[illegible]

```

Alignment Scores:
Pred. No.:      859      Length:      2714
Score:          41.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             8       Gaps:        0

US-10-774-176-8 (1-9) x AB168308 (1-2714)

QY      1 Ala1lePheLeuLeuValLeuTyrLeu 9
Db      1865 GCTATTTCCTCCTCGTTGTTGATTG 1891

RESULT 34
HSA012159      HSA012159      5551 bp      DNA      linear      PRI 15-APR-2005
LOCUS          Homo sapiens 574 oncofetal trophoblast glycoprotein gene.
DEFINITION     AJ012159
ACCESSION      AJ012159
VERSION        AJ012159.1 GI:3805946
KEYWORDS       574 gene; 574 oncofetal trophoblast glycoprotein.
SOURCE         Homo sapiens
ORGANISM       Homo sapiens
               Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Hominidae; Homo.
REFERENCE      1
AUTHORS        King, K.W., Sheppard, P.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE          Organisation of the mouse and human 574 oncofoetal leucine-rich
               glycoprotein genes and expression in foetal and adult murine
               tissues
JOURNAL        Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED         10366710
REFERENCE      2 (bases 1 to 5551)
AUTHORS        Myers, K.A.
TITLE          Direct Submission
JOURNAL        Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
               Institute for Cancer Research, Christie Hospital, Wilmslow Road,
               Manchester, M20 9BX, UK
FEATURES       Location/Qualifiers
source         1..5551
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               2698..2703
               /bound_moiety="Sp1"
               2704..2709
               /bound_moiety="Sp1"
               2716..2740
               /gene="574"
               2716..2800
               /gene="574"
               /evidence=experimental
               2801..3092
               /gene="574"
               /evidence=experimental
               3093..5400
               /gene="574"
               /evidence=experimental
               3431..4693
               /gene="574"
               /codon_start=1
               /product="574 oncofetal trophoblast glycoprotein"
               /protein_id="CAA09930.1"
               /db_xref="GI:3805947"
               /db_xref="GOA:Q13641"
               /db_xref="InterPro:IPR000372"
               /db_xref="InterPro:IPR000483"
               /db_xref="InterPro:IPR001611"
               /db_xref="InterPro:IPR003591"
               /db_xref="UniProt/TREMBL:Q13641"
               /translation="WFGCSRGPAAGDGRRLRLARLALVLLGWSSSSPTSSASSFSSS
               APFLASAVSAQPLDQCPCSEARTKCVNRNLTEVPTDLPAYVRNFLTNGQ

LAVLPAGAFARRPPLAEALANLNGSRLDEVRAFAEHLPSLRQLDLSHNPLADLSPF
AFSGSNASVAPSPLVELIINHIVPPEDQRNRSFEGMVVAALLAGRALQGRLELA
SNHELVLPRDLVAQLPSLRHLDLSNNLSVSLTVYSPRNLTHLSLHLEDNALKVLHNG
TLAELOGLPHIRVFLDNNPWCDCHWADMTWLKETEYVQGDRLTTCAYPEKWRVRL
LELNSADLDCDPLPSPSLQTSYVFLGIVLALIGAIFLLVLYLNKRGIGKKWNNIRDAC
RDHMEGYHYRYEINADPRLTNLSNSDV"
sig_peptide   3431..3516
mat_peptide   /gene="574"
              3517..4690
              /gene="574"
polyA_signal   /product="574 oncofetal trophoblast glycoprotein"
              5331..5336
              /gene="574"
polyA_signal   5380..5385
              /gene="574"

ORIGIN
Alignment Scores:
Pred. No.:      1.55e+03      Length:      5551
Score:          41.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             8       Gaps:        0

US-10-774-176-8 (1-9) x HSA012159 (1-5551)

QY      1 Ala1lePheLeuLeuValLeuTyrLeu 9
Db      4532 GCTATTTCCTCCTCGTTGTTGATTG 4558

RESULT 35
MMU012160      MMU012160      7942 bp      DNA      linear      ROD 15-APR-2005
LOCUS          Mus musculus 574 oncofetal trophoblast glycoprotein gene.
DEFINITION     AJ012160
ACCESSION      AJ012160
VERSION        AJ012160.1 GI:3805948
KEYWORDS       574 gene; 574 oncofetal trophoblast glycoprotein.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
               Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1
AUTHORS        King, K.W., Sheppard, P.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE          Organisation of the mouse and human 574 oncofoetal leucine-rich
               glycoprotein genes and expression in foetal and adult murine
               tissues
JOURNAL        Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED         10366710
REFERENCE      2 (bases 1 to 7942)
AUTHORS        Myers, K.A.
TITLE          Direct Submission
JOURNAL        Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
               Institute for Cancer Research, Christie Hospital, Wilmslow Road,
               Manchester, M20 9BX, UK
FEATURES       Location/Qualifiers
source         1..7942
               /organism="Mus musculus"
               /mol_type="genomic DNA"
               /strain="129/Sv"
               /db_xref="taxon:10090"
               /clone_lib="Lambda Dash"
               3108..3113
               /bound_moiety="Sp1"
               3114..3119
               /bound_moiety="Sp1"
               3124..5779
               /gene="574"
               3124..3151
               /gene="574"
               3152..3450
               /gene="574"
               intron

```

exon 3451..5779
/gene="5T4"
CDS 3779..5059
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA09931.1"
/db_xref="GI:3805949"
/db_xref="GOA:Q9Z0L0"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProt/TTrEMBL:Q9Z0L0"
/translation="MPGAGSGPSAGDGRRLRLRLALVLGHWVSASAPSSVPSSSTS
PADFLASGAPPPAERCPAAECESHAARTVKVNRNLLLEVPADLPFYVRNLIPLTGNQ
MTVLPAAGAPQPLADLEALNLGNHLKEVCAGAFELPGLRLRLDLSHNPLTNLSAF
VFAGSNASVSPLEELILNHIIVPEQRQNGSPFGMVAFEGVAAALRGLALRGL
TRLASNHFTPLPRDLAQLPSLRYLRLNNSLVSLTYASFRNLTHLSLHLEDNAL
KVLNHTLAEGHGVKVLDDNPMVCDYMDVMWLKETEVPDKARLTCAPEK
MRNGLLDLSDDLDCDAVLQSLQTSYVFLGIVLALIGALFLVLVLYLRKGIKKMMH
NIRDACRDMHGYRYRYBINADPRLTNLSNSDV"
sig_peptide 3779..3865
/gene="5T4"
mat_peptide 3866..5056
/gene="5T4"
polyA_signal /product="5T4 oncofetal trophoblast glycoprotein"
5713..5718
/gene="5T4"
polyA_signal 5759..5764
/gene="5T4"

ORIGIN

Alignment Scores: 2.08e+03 Length: 7942
Pred. No.: 41.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 9

US-10-774-176-8 (1-9) x MMU012160 (1-7942)

Oy 1 AlallePheLeuLeuValleuTyrieu 9
Db 4898 GCTATTTCCTCCTCGTGTGTTATTTG 4924

RESULT 36
AP008208_357/c
WPCOMMENT
Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

Fragment Name	Begin	End
AP008208_000	1	110000
AP008208_001	100001	210000
AP008208_002	200001	310000
AP008208_003	300001	410000
AP008208_004	400001	510000
AP008208_005	500001	610000
AP008208_006	600001	710000
AP008208_007	700001	810000
AP008208_008	800001	910000
AP008208_009	900001	1010000
AP008208_010	1000001	1110000
AP008208_011	1100001	1210000
AP008208_012	1200001	1310000
AP008208_013	1300001	1410000
AP008208_014	1400001	1510000
AP008208_015	1500001	1610000
AP008208_016	1600001	1710000
AP008208_017	1700001	1810000
AP008208_018	1800001	1910000
AP008208_019	1900001	2010000

AP008208_020 2110000
AP008208_021 2210000
AP008208_022 2310000
AP008208_023 2410000
AP008208_024 2510000
AP008208_025 2610000
AP008208_026 2710000
AP008208_027 2800001
AP008208_028 2910000
AP008208_029 3000001
AP008208_030 3110000
AP008208_031 3210000
AP008208_032 3300001
AP008208_033 3400001
AP008208_034 3510000
AP008208_035 3600001
AP008208_036 3700001
AP008208_037 3800001
AP008208_038 3900001
AP008208_039 4010000
AP008208_040 4110000
AP008208_041 4200001
AP008208_042 4310000
AP008208_043 4400001
AP008208_044 4510000
AP008208_045 4600001
AP008208_046 4700001
AP008208_047 4810000
AP008208_048 4900001
AP008208_049 5010000
AP008208_050 5110000
AP008208_051 5210000
AP008208_052 5310000
AP008208_053 5410000
AP008208_054 5510000
AP008208_055 5600001
AP008208_056 5700001
AP008208_057 5810000
AP008208_058 5900001
AP008208_059 6010000
AP008208_060 6100001
AP008208_061 6210000
AP008208_062 6300001
AP008208_063 6410000
AP008208_064 6500001
AP008208_065 6600001
AP008208_066 6700001
AP008208_067 6810000
AP008208_068 6900001
AP008208_069 7000001
AP008208_070 7110000
AP008208_071 7210000
AP008208_072 7300001
AP008208_073 7410000
AP008208_074 7500001
AP008208_075 7600001
AP008208_076 7700001
AP008208_077 7800001
AP008208_078 7900001
AP008208_079 8000001
AP008208_080 8110000
AP008208_081 8210000
AP008208_082 8310000
AP008208_083 8410000
AP008208_084 8510000
AP008208_085 8600001
AP008208_086 8700001
AP008208_087 8800001
AP008208_088 8900001
AP008208_089 9010000
AP008208_090 9110000
AP008208_091 9210000
AP008208_092 9310000

AP008208_093 9410000
AP008208_094 9510000
AP008208_095 9610000
AP008208_096 9700001
AP008208_097 9810000
AP008208_098 9900001
AP008208_099 10010000
AP008208_100 10110000
AP008208_101 10210000
AP008208_102 10310000
AP008208_103 10410000
AP008208_104 10500001
AP008208_105 10610000
AP008208_106 10710000
AP008208_107 10810000
AP008208_108 10910000
AP008208_109 11010000
AP008208_110 11100001
AP008208_111 11210000
AP008208_112 11310000
AP008208_113 11400001
AP008208_114 11510000
AP008208_115 11610000
AP008208_116 11710000
AP008208_117 11810000
AP008208_118 11910000
AP008208_119 12010000
AP008208_120 12110000
AP008208_121 12210000
AP008208_122 12300001
AP008208_123 12410000
AP008208_124 12510000
AP008208_125 12610000
AP008208_126 12710000
AP008208_127 12810000
AP008208_128 12910000
AP008208_129 13010000
AP008208_130 13100001
AP008208_131 13210000
AP008208_132 13310000
AP008208_133 13410000
AP008208_134 13510000
AP008208_135 13610000
AP008208_136 13710000
AP008208_137 13810000
AP008208_138 13910000
AP008208_139 14010000
AP008208_140 14110000
AP008208_141 14210000
AP008208_142 14310000
AP008208_143 14410000
AP008208_144 14510000
AP008208_145 14610000
AP008208_146 14710000
AP008208_147 14810000
AP008208_148 14910000
AP008208_149 15010000
AP008208_150 15110000
AP008208_151 15210000
AP008208_152 15310000
AP008208_153 15410000
AP008208_154 15510000
AP008208_155 15610000
AP008208_156 15710000
AP008208_157 15810000
AP008208_158 15910000
AP008208_159 16010000
AP008208_160 16110000
AP008208_161 16210000
AP008208_162 16310000
AP008208_163 16410000
AP008208_164 16510000
AP008208_165 16610000

AP008208_166 16710000
AP008208_167 16810000
AP008208_168 16910000
AP008208_169 17010000
AP008208_170 17110000
AP008208_171 17210000
AP008208_172 17310000
AP008208_173 17410000
AP008208_174 17510000
AP008208_175 17610000
AP008208_176 17710000
AP008208_177 17810000
AP008208_178 17910000
AP008208_179 18010000
AP008208_180 18110000
AP008208_181 18210000
AP008208_182 18310000
AP008208_183 18410000
AP008208_184 18510000
AP008208_185 18610000
AP008208_186 18710000
AP008208_187 18810000
AP008208_188 18910000
AP008208_189 19010000
AP008208_190 19110000
AP008208_191 19210000
AP008208_192 19310000
AP008208_193 19410000
AP008208_194 19510000
AP008208_195 19610000
AP008208_196 19710000

Alignment Scores:

Pred. No.:	1.82e+04	Length:	110000
Score:	41.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	15	Gaps:	0

US-10-774-176-8 (1-9) x AP008208_357 (1-110000)

QY 1 AlAlaPheLeuLeuValLeuTyrLeu 9
|||
Db 82807 GCAATATTCTGTAGTCTTATTTA 82781

RESULT 37

AP004082/c

LOCUS AP004082 120743 bp DNA linear PLN 15-JUN-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone:QJ1149_C12.

ACCESSION AP004082 GI:48716388

KEYWORDS

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

1 Sasaki, T., Matsumoto, T. and Yamamoto, K.

Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC

clone:QJ1149_C12

Published Only in Database (2001)

2 (bases 1 to 120743)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission

Submitted (22-AUG-2001) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/.

Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT

On Jun 14, 2004 this sequence version replaced gi:38142420.

Genes were predicted from the integrated results of the following:
 GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>), FGENESH
 (<http://www.softberry.com/>), GeneMark.hmm
 (<http://opal.biology.gatech.edu/GeneMark/>), GlimmerM
 (http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM
 (<http://rgp.dna.affrc.go.jp/RiceHMM/>), SplicePredictor
 (<http://bioinformatics.iastate.edu/cgi-bin/sp.cgi>), sim4
 (<http://globin.cse.psu.edu/html/docs/sim4.html>), gap2
 (<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI Nonredundant Protein
 database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA
 sequence database at RGP or DDBJ. Protein homologues of the coding
 regions were searched against NCBI Nonredundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from -21M13 to M13rev of the BAC
 clone. This sequence of OJ1149 C12 clone has an overlap with
 OJ1124_D06 (DDBJ: AP004043) clone at 5' end and with OJ1282_E10
 (DDBJ: AP005290) at 3' end. The sequence was generated by combining
 Monsanto and RGP-Japan sequencing data. Detailed information on
 overlap and assembly quality together with annotation of this entry
 is available at
<http://rgp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES

source

Location/Qualifiers
 1..120743

/organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1149_C12"
 /complement(2011..4039)
 /gene="OJ1149_C12.1"
 /complement(join(<2011..2166,2833..2916,3398..3489,
 4009..>4039))
 /gene="OJ1149_C12.1"
 /note="start and end point are not identified"
 /complement(join(2011..2166,2833..2916,3398..3489,
 4009..4039))
 /gene="OJ1149_C12.1"
 /note="predicted by FGENESH etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD22998.1"
 /db_xref="GI:48716389"
 /translation="MINLTSSVSLWMDKNKVALGRNYSLLIVPLILLQIDP
 QSSSCWATVAFIALNRYLLSLAPQESACINGAIGDAVAQSYLERIGDRSGE
 KVELEKKAISGFTCKPG"
 /complement(5014..5134)
 /gene="OJ1149_C12.5"
 /complement(5014..5134)
 /gene="OJ1149_C12.5"
 /note="3' terminal repeat"
 /rpt_type="terminal"
 /complement(join(6772..6816,6937..6971,7024..7099))
 /gene="OJ1149_C12.2"
 /complement(join(6772..6816,6937..6971,7024..7099))
 /gene="OJ1149_C12.2"
 /note="hypothetical ORF
 predicted by GlimmerM

gene
 mRNA
 CDS
 gene
 misc_feature
 gene
 misc_feature
 gene
 misc_feature
 gene
 mRNA
 CDS
 gene
 repeat_unit
 gene
 mRNA
 CDS
 gene
 repeat_unit
 gene
 misc_feature

this category is not included in IRGSP standard"
 18095..19752
 /gene="OJ1149_C12.3"
 18095..19752_
 /gene="OJ1149_C12.3"
 /note="supported by full-length cDNA(s): AK060482"
 18180..19226
 /gene="OJ1149_C12.3"
 /note="contains full-length cDNA(s): AK060482"
 /codon_start=1
 /product="unknown protein"
 /protein_id="BAD22999.1"
 /db_xref="GI:48716390"
 /translation="MPDEFIDNGYTSDLHRLSRTTGERLLSVTETGKVPFEGVQP
 LVKLMSVSHGSAIEVLRSITPTCDISFGSPETRLIMHTTGLLPMGLQLTQDAS
 SLGSSPQEQNQKAYVNASNIWCRVKSLLDLAEVFRAYSFGIISLELDLFAASFP
 PQSSFPKSHSLAPFGLHLLRLRSGFDYQRVVLLMKSLLOQTTPVDPSPQIPQVNVV
 SIVSALCAEALNVLRLRSCVGTGGDDIGFGNGHGMKGVHQSMLLPOSSP
 KARSPGLQYAAAGSGFTLMGQGGGAADTVATRDVALQNTLLRLGRVLDTCALGRK
 RDHRLVFPVANIG"
 /complement(join(20355..20681,20760..21011))
 /gene="OJ1149_C12.4"
 /complement(join(20355..20681,20760..21011))
 /gene="OJ1149_C12.4"
 /note="hypothetical ORF
 predicted by GENSCAN
 this category is not included in IRGSP standard"
 /complement(24066..27050)
 /gene="OJ1149_C12.5"
 /complement(24066..27050)
 /gene="OJ1149_C12.5"
 /note="mutator-like transposase-like"
 /complement(27264..27434)
 /gene="OJ1149_C12.6"
 /complement(27264..27434)
 /gene="OJ1149_C12.6"
 /note="hypothetical ORF
 predicted by GlimmerM
 this category is not included in IRGSP standard"
 /complement(27727..28448)
 /gene="OJ1149_C12.7"
 /complement(join(<27727..27781,28261..>28448))
 /gene="OJ1149_C12.7"
 /note="start and end point are not identified"
 /complement(join(27727..27781,28261..28448))
 /gene="OJ1149_C12.7"
 /note="predicted by RiceHMM etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD23000.1"
 /db_xref="GI:48716391"
 /translation="MDEKKERRRKVKERQDLAQGGGEVDVAGAPAAAGFFPGVDG
 RSHGVHGMRCGMDGFSAYLSANRFGAEPVNLAK"
 /complement(29172..29291)
 /gene="OJ1149_C12.5"
 /complement(29172..29291)
 /gene="OJ1149_C12.5"
 /note="5' terminal repeat"
 /rpt_type="terminal"
 30234..31789
 /gene="OJ1149_C12.8"
 30234..31789
 /gene="OJ1149_C12.8"
 /note="supported by full-length cDNA(s): AK060600"
 30316..31404
 /gene="OJ1149_C12.8"
 /note="contains EST(s): D47685(S13323)
 contains full-length cDNA(s): AK060600
 similar to Oryza sativa chromosome 4, OSJNB0004A17.4"
 /codon_start=1
 /product="unknown protein"
 /protein_id="BAD23001.1"
 /db_xref="GI:48716392"

```

/translation="MLRQTHDSYGPERSHRGSAEKLSESRFSOREPMPMLTADSSDLR
CELLKORLNPRSQSDRHNVPEDRDRHQREQASNDGVSRLRGLRLP
ATTTFRLGLQPEKRDGPRALRPPSQTDLRKLDKAKPDRVSGNVQSSLSK
ANEDASLNFAGPKFLAEILKAKVAGSLMKSSRLTGPVMTSEIVTIKSSDPVLFD
GPKPLNALIKRREASDNATFGKREHSGDSEGSQNDFRNEDDITVGNWTEGNG
BEAFQEDVVVYDLSLSPADIAAABADDAARELSEQQDVETAEYDEMDDVNAARE
NDQRYEBDEDDLEDDDPARKVGMIT"
complement(32006..33785)
/gene="Q11149_C12.9"
complement(join(32006..32537,33266..33785))
/gene="Q11149_C12.9"
/note="supported by full-length cDNA(s): AK098955"
complement(join(32120..32537,33266..33783))
/gene="Q11149_C12.9"
/note="contains full-length cDNA(s): AK061847, AK098955
similar to Oryza sativa chromosome 4, OSJNB0085F13.19"
/codon_start=1
/product="unknown protein"
/protein_id="BAD23002.1"
/db_xref="GI:48716393"
/translation="MIGFIGGVAGALPLVLRTRTCTPPPPMAMTPLHHTCSFFPRP
QPSLSHSPCAAADDDFTVDYDPEEKEDEBGSWEGAVVYRRDASVHLRYAT
TLERGLGLSGSTRARAATGILISLTNTGKDDTPVLVSLDVARRGRRLRD
GIVRTVITLCYCAEPAPQGI FANFSILLTDRVEEPDVLDGTIPREKTKAPSLT
GSQEGDDEDIDWDRLHPAGEKSIDISKHIDIHLITLDALCSPCKGLCGCG
ENLNTSSCSNAKQQAQNVQRGPKLKKLPQR"
complement(36443..40501)
/gene="Q11149_C12.10"
complement(join(36443..36783,36987..37068,37163..37200,

```

Alignment Scores:

```

Pred. No.: 1.97e+04 Length: 120743
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

```

US-10-774-176-8 (1-9) x AP004082 (1-120743)

```

Oy 1 AlaliePheLeuLeuValleuTyrlieu 9
Db 53522 GCATATCTCTGTTAGTCTTTATTTA 53496

```

```

RESULT 38
HSJ492P14 121909 bp DNA linear PRI 18-MAY-2005
LOCUS HSJ492P14
DEFINITION Human DNA sequence from clone RP3-492P14 on chromosome 6q13-15
Contains a single stranded DNA binding protein pseudogene, the TPBG
gene for trophoblast glycoprotein (574-AG) and a CpG island,
complete sequence.
ACCESSION AL121977
VERSION AL121977.11 GI:11863678
KEYWORDS HTG; CpG island; TPBG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 121909)
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Dec 15, 2000 this sequence version replaced gi:11558491.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., ENBL; Sw., SWISSPROT; Tr., TRMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping

```

Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>
RP3-492P14 is from the library RPCI-3 constructed by the group of
Piet de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES

```

Location/Qualifiers
source
1..121909
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosomes="6"
/map="q13-15"
/clone="RP3-492P14"
/clone_lib="RPCI-3"
100
/note="Clone right end: RP1-93K22"
complement(10004..10982)
/locus_tag="RP3-492P14.2-001"
/pseudo
CDS
complement(10004..10982)
/locus_tag="RP3-492P14.2-001"
/note="match: proteins: P81877 Q99LX9 Q9BWW6 Q9CYZ8 Q9D6L4
Q9P038 Q9Y477"
/misc_feature
/codon_start=1
86539
/note="Clone left end: RP1-90G1"
109639..118636
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
join(109639..109916,110631..116836)
/mRNA
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
/product="trophoblast glycoprotein"
/note="match: ESTs: AA149121 AA152323 AA565852 AA643734
AL544610 AW471072 AW662538 BB260089 BF306457 BF306926
BF314984 B1196133 B1562387 BM069633 BM670613
match: cDNAs: AJ420536.1 Z29083.1"
110970..112232
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
/standard_name="OTTHUMP0000016786"
/note="match: proteins: Q13641 Q9QYD9 Q9Z0L0"
/codon_start=1
/product="trophoblast glycoprotein"
/protein_id="CAI21546.1"
/db_xref="GI:56203539"
/db_xref="Genbank:12004"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TrEMBL:Q13641"
/translation="MPGSGRGPAGDGRLLARLALVLLGWSSSPSSPTSSASSFSSS
APFLASVQAQPLPDPCCPALCESSAARTVKVNRNLTEVPTDLPAYVNRNLFTGSP
LAVLPAGAPARPLPBLAALNLGSRSLRDEVRAGAFHPLFSQLDLSHNPDLSTQF
AFSGSNASVAPSPLVELILNIHVPPEQRQRNPFEGMVVAALLAGRALQGLRLELA
SNHFLYLPDVLALQLPSLRHLDSLNNLSVSLTVSPRNLTHLSLHLEDNALVKVLHNG

```



```

/product="5'-nucleotidase, ecto (CD73)"
/protein_id="CAI40167.1"
/db_xref="GI:57162422"
/db_xref="GOA:Q96B60"
/db_xref="InterPro:IPR004843"
/db_xref="InterPro:IPR006146"
/db_xref="InterPro:IPR006179"
/db_xref="UniProt/TREMBL:Q96B60"
/translation="MCPRAARAPATVILLALGAVLWPAAGAWELTILHTNDVHSRLQET
SDSSKCNASCMGVKARLFTKQVQIRABENVLLDAGDQYQGTITFTVYKGAEVA
YKVLAVDAMALGNHFDNGVEGLTEPLLKEAKFPILSANIKAKGPKLASQISGLYLP
FFMLPVGVGVIGVGTSTPTPLNPGTNLVEDEITAIQPEVDFKLTNLVKNKILAL
GHSGFEMDKLIAQKVRGVDDVVVGHSNTFLYTGNCFKRIAWARMS"
join(<15693, 15737, 29547, 29744, 31647, 31801, 33806, 33911,
36289, >36473)
/gene="NT58"
/locus_tag="RP11-321N4.1-003"
/locus_tag="RP11-321N4.1-003"
/product="5'-nucleotidase, ecto (CD73)"
/note="match: ESTs: BG291434"
join(<15693, 15737, 29547, 29744, 31647, 31801, 33806, 33911,
36289, >36473)
/gene="NT58"
/locus_tag="RP11-321N4.1-003"
/locus_tag="RP11-321N4.1-003"
/standard_name="OTTHUMP00000016810"
/codon_start=3
/product="5'-nucleotidase, ecto (CD73)"
/protein_id="CAI40164.1"
/db_xref="GI:57162419"
/db_xref="GOA:Q5JRQ2"
/db_xref="UniProt/TREMBL:Q5JRQ2"
/translation="DVVVGHSNTFLYTGNPSPKVPAGKVPPIVTSDDGRKVPVVOA
YAFQVGLVLTPEPGRNVISSHGNPILNSSIPEDPSIKADINKWRILKLNYSVTOE
LQKTVLDGSSQSCRPECNMGNLICDAMINNNLRHTDEMFWNHSVCMILNGGGRS
PIDERNGHVVYDLSRKPGRVVKLVLCVTKRVPSPYDPLKQDEVKYVILNPLANG
GDGFQMIKDEL"
15881, 15886
/gene="NT58"
/locus_tag="RP11-321N4.1-002"
15908
/gene="NT58"
/locus_tag="RP11-321N4.1-002"
join(<29709, 29744, 31647, 31801, 33806, 33911, 34820, 34969,
35347, 35637)
/gene="NT58"
/locus_tag="RP11-321N4.1-004"
join(<29709, 29744, 31647, 31801, 33806, 33911, 34820, 34969,
35347, 35637)
/gene="NT58"
/product="5'-nucleotidase, ecto (CD73)"
/note="match: ESTs: A1580514 BQ001980"
join(<29709, 29744, 31647, 31801, 33806, 33911, 34820, 34969,
35347, 35438)
/gene="NT58"
/locus_tag="RP11-321N4.1-004"
/standard_name="OTTHUMP00000016811"
/codon_start=3
/product="5'-nucleotidase, ecto (CD73)"
/protein_id="CAI40165.1"
/db_xref="GI:57162420"
/db_xref="GOA:Q5JRQ1"
/db_xref="UniProt/TREMBL:Q5JRQ1"
/translation="NPILNNSIPEDPSIKADINKWRILKLNYSVTOELGKTVILDGSS
SQSQRFCNMGNLICDAMINNNLRHTDEMFWNHSVCMILNGGGRSPIDERNNGTIT
WENLAALVPGGTFTDLVOLKGSITLKAPHSVHRYQSGSTGEFLQVGDLEKCCPICINQ
DDSWKEQTAQYILNTWF"
40076, 40081
/gene="NT58"
/locus_tag="RP11-321N4.1-001"

```

```

gene
join(complement(AL136082..22:8981, .9287),
complement(118570, .118690), complement(1116610, .116686),
complement(116449, .116527), complement(102288, .102372),
complement(94035, .94191), complement(92613, .92688),
complement(91818, .91862), complement(91629, .91709),
complement(91424, .91538), complement(87917, .88072),
complement(87494, .87618), complement(86297, .86355),
complement(81104, .81236), complement(77911, .78112),
complement(74504, .74587), complement(72574, .72674),
complement(70438, .70549), complement(62320, .62360),
complement(62360, .62360))

Alignment Scores:
Pred. No.: 2, 08e+04 Length: 129010
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-8 (1-9) x AL589666 (1-129010)

QY 1 AlalleleLeuLeuValleyrleu 9
Db 103534 GCGATATTTTGGCTAGTATTATACCTC 103560

RESULT 40
AC158516 167046 bp DNA linear ROD 21-JUN-2005
LOCUS Mus musculus BAC clone RP24-511A23 from chromosome 9, complete
DEFINITION
sequence.
ACCESSION AC158516 AC117768
VERSION AC158516.2 GI:63025421
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 167046)
AUTHORS Adams,S., Cotton,M. and Haglund,K.
TITLE The sequence of Mus musculus BAC clone RP24-511A23
JOURNAL Unpublished (2001)
AUTHORS Wilson,R.K.
REFERENCE
2 (bases 1 to 167046)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
3 (bases 1 to 167046)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
4 (bases 1 to 167046)
AUTHORS Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2005) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On May 4, 2005 this sequence version replaced gi:61656412.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submission@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0511A23
Drafting center: WIBR
NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e. phred quality

```

>=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. The assembly was confirmed by restriction digest. This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal.

If a sequence meets the criteria of the above statement, it needs no comments or tags. If the criteria are not met, such as ambiguous bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

FEATURES

```

source
    Location/Qualifiers
        1..167046
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="9"
            /clone_lib="RPCI-24"
            /clone="RP24-511A23"
            16685..16712
                /note="Sequence derived from PCR product of genomic DNA"
            31565..31779
                /note="Unresolved simple sequence repeat."
            46721..46808
                /note="Unresolved simple sequence repeat."
            142336..142347
                /note="Sequence derived from one plasmid subclone."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      2.57e+04      Length:      167046
Score:          41.00        Matches:      9
Percent Similarity: 100.0%    Conservative: 0
Best Local Similarity: 100.0% Mismatches:      0
Query Match:      100.0%     Indels:         0
DB:              9          Gaps:           0

```

US-10-774-176-8 (1-9) x AC158516 (1-167046)

Oy 1 AlallePheLeuValleyrYleu 9

Db 109719 GCTATTTCCTCCTGTTTGTATTG 109693

RESULT 41

```

LOCUS      AC155741/c
DEFINITION Bos taurus clone CH240-40N24, WORKING DRAFT SEQUENCE, 13 unordered
            pieces
ACCESSION  AC155741
VERSION    AC155741.2 GI:68265036
KEYWORDS   HTGS; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;

```

Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 208985)

REFERENCES

AUTHORS

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, K., Garcia, A., Garner, T., Garza, M., Gregorogis, E., Geer, K., Gill, R., Grady, M., Guerra, T., Guvera, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louisaeged, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martine, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwakoelameh, O., Okwunog, G., Olarnpungsoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plapper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Satter, S., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlecyk, R., Woodden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

TITLE

JOURNAL

Unpublished

2 (bases 1 to 208985)

Worley, K.C.

Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 208985)
 Cow Genome Sequencing Consortium.
 Direct Submission
 Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Jun 28, 2005 this sequence version replaced gi:57900766.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FCLO
Center clone name: CH240-40N24

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 203744 bases at least Q40

Consensus quality: 205219 bases at least Q30

Consensus quality: 206413 bases at least Q20

Estimated insert size: 209875; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 20170: contig of 20170 bp in length
* 20171 20220: gap of 50 bp
* 20221 26553: contig of 6333 bp in length
* 26554 26603: gap of 50 bp
* 26604 66878: contig of 40275 bp in length
* 66879 66928: gap of 50 bp
* 66929 69265: contig of 2337 bp in length
* 69266 69315: gap of 50 bp
* 69316 98544: contig of 29229 bp in length
* 98545 98594: gap of 50 bp
* 98595 109039: contig of 10445 bp in length
* 109040 109089: gap of 50 bp
* 109090 171503: contig of 62414 bp in length
* 171504 171603: gap of unknown length
* 171604 178456: contig of 6853 bp in length
* 178457 178506: gap of 50 bp
* 178507 182908: contig of 4402 bp in length
* 182909 182958: gap of 50 bp
* 182959 198621: contig of 15662 bp in length
* 198622 198740: gap of 120 bp
* 198741 205490: contig of 6750 bp in length
* 205491 205590: gap of unknown length
* 205591 206826: contig of 1236 bp in length
* 206827 206926: gap of unknown length
* 206927 208985: contig of 2059 bp in length.

FEATURES

Location/Qualifiers
1. 208985
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-40N24"
20171..20220
/estimated_length=50
26554..26603
/estimated_length=50
66879..66928
/estimated_length=50
69266..69315
/estimated_length=50
98545..98594
/estimated_length=50
109040..109089

/estimated_length=50
171504..171603
/estimated_length=unknown
178457..178506
/estimated_length=50
182909..182958
/estimated_length=50
198621..198740
/estimated_length=120
205491..205590
/estimated_length=unknown
206827..206926
/estimated_length=unknown

ORIGIN

Alignment Scores:
Pred. No.: 3.1e+04 Length: 208985
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-8 (1-9) x AC155741 (1-208985)

Qy 1 AlAllePhelLeuValLeuTyLeu 9
|||||

Db 33653 GCCATTTTCCTCTCTGCTGTACCTG 33627

RESULT 42

AC128294/c

LOCUS

DEFINITION

AC128294

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC128294 210237 bp DNA linear HTG 19-NOV-2002
Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.
AC128294
AC128294.3 GI:25083347
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 210237)
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angulano, D.,
Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, S.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guvarra, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S., Hlyk, S., Hume, J., Idlebird, D., Jackson, A.,
Hollins, B., Howells, S., Hulyk, S., Jang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Khan, Z., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhera, L., Loulseghe, H., Lozano, R., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M., McNeill, T., Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plapper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabak, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 210237)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 210237)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23265004.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZGV
Center clone name: CH230-176H20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 201781 bases at least Q40
Consensus quality: 203921 bases at least Q30
Consensus quality: 205310 bases at least Q20
Estimated insert size: 205531; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced as soon as it is available and by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 210237: contig of 210237 bp in length.

Location/Qualifiers
1..210237
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-176H20"
1..1142
/note="wgs end extension
clone_end:17"
2177..144799
/note="clone boundary
clone_end:17"
site:
end sequence: BH360464"
complement(206062..206961)
/note="clone boundary
clone_end:Sp6
site:
end sequence: BH360465"
208907..210237
/note="wgs end extension
clone_end:Sp6"
FEATURES
source
misc_feature
misc_feature
misc_feature
misc_feature
ORIGIN
Alignment Scores:
Pred. No.: 3.11e+04 Length: 210237
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0
US-10-774-176-8 (1-9) x AC128294 (1-210237)
QY 1 AlaiIlePheLeuLeuValuYrLeu 9
|||||
Db 110399 GCTATTTCCTCTCGTTTGATTG 110373
RESULT 43
AC156733/c
LOCUS
DEFINITION
AC156733.2 GI:68265384
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Bos taurus clone CH240-40B14, *** SEQUENCING IN PROGRESS ***, 20
unordered pieces.
AC156733
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bialwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, P., Garcia, A., Garner, T., Garza, M., Guebara, P., Haaland, W., Hamill, C., Hamilton, C., Guerri, M., Guevara, W., Gunaratne, P., Geer, K., Gill, R., Grady, M., Grady, M., Guzman, R., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowls,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensheewa,L., Loudeged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Mallory,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,P., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puozo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savory,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartebeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,D., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasaana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P., Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 218269)
Worley,K.C.

Direct Submission
Submitted (03-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 218269)

Cow Genome Sequencing Consortium.
Direct Submission
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 28, 2005 this sequence version replaced gi:58531453.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FCOA
Center clone name: CH240-40B14
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 211442 bases at least Q40
Consensus quality: 213025 bases at least Q30
Consensus quality: 214500 bases at least Q20
Estimated insert size: 217823; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 20 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 47698: contig of 47698 bp in length
47699 47748: gap of 50 bp
47749 53619: contig of 5871 bp in length
53620 53669: gap of 50 bp
53670 58241: contig of 4572 bp in length
58242 58291: gap of 50 bp
58292 65034: contig of 6763 bp in length
65035 65104: gap of 50 bp
65105 71554: contig of 6450 bp in length
71555 71555: gap of unknown length
71556 101839: contig of 30184 bp in length
101840 101889: gap of 50 bp
101890 119077: contig of 17189 bp in length
119078 119127: gap of 50 bp
119128 121183: contig of 2056 bp in length
121184 121233: gap of 50 bp
121234 128781: contig of 7548 bp in length
128782 128831: gap of 50 bp
128832 137035: contig of 8204 bp in length
137036 137085: gap of 50 bp
137086 145862: contig of 8777 bp in length
145863 145912: gap of 50 bp
145913 161254: contig of 15342 bp in length
161255 161304: gap of 50 bp
161305 191171: contig of 29867 bp in length
191172 191221: gap of 50 bp
191222 192340: contig of 1119 bp in length
192341 192390: gap of 50 bp
192391 206481: contig of 14091 bp in length
206482 206531: gap of 50 bp
206532 210457: contig of 3925 bp in length
210458 210556: gap of unknown length
210557 211589: contig of 1033 bp in length
211590 211689: gap of unknown length
211690 213759: contig of 2070 bp in length
213760 213859: gap of unknown length
213860 215381: contig of 1522 bp in length
215382 215481: gap of unknown length
215482 218269: contig of 2788 bp in length.

FEATURES
Location/Qualifiers
1. 218269
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-40B14"
47699. 47748
/estimated_length=50
53620. 53669
/estimated_length=50
58242. 58291
/estimated_length=50
65055. 65104
/estimated_length=50
71555. 71654
/estimated_length=unknown
101839. 101888
/estimated_length=50
119078. 119127
/estimated_length=50
121184. 121233
/estimated_length=50
128782. 128831
/estimated_length=50
137036. 137085
/estimated_length=50

```

gap      145863..145912
/estimated_length=50
gap      161255..161304
/estimated_length=50
gap      191172..191221
/estimated_length=50
gap      192341..192390
/estimated_length=50
gap      206482..206531
/estimated_length=50
gap      210457..210556
/estimated_length=unknown
gap      211590..211689
/estimated_length=unknown
gap      213760..213859
/estimated_length=unknown
gap      215382..215481
/estimated_length=unknown

ORIGIN
Alignment Scores:
Pred. NO.:      3.21e+04      Length:      218269
Score:          41.00        Matches:      9
Percent Similarity: 100.0%    Conservative: 0
Best Local Similarity: 100.0% Mismatches:      0
Query Match:    100.0%      Indels:        0
DB:             14          Gaps:          0

US-10-774-176-8 (1-9) x AC156733 (1-218269)
Qy      1 AlattlePheLeuValLeuTyrieu 9
Db      5727 GCCATTTCCTCTGTGCTGTGACCTG 5701

RESULT 44
AC106962/C
LOCUS      AC106962
DEFINITION Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4
unorderd pieces.
ACCESSION AC106962
VERSION   AC106962.5 GI:25139469
KEYWORDS  HTGS_PHASE1, HTGS_DRAFT, HTGS_FULLTOP.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Rattus.
1 (bases 1 to 239076)
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alabrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,P.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgescu,E., Gear,K., Gill,R., Grady,M., Guerra,M., Guevara,M.,
Gunnarsson,P., Healand,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Haviak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louisseged,H., Lozado,R.J., Lu,X., Ma,J.,

```

```

Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Naik,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpungoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,B., Song,X.-Z., Sorelle,R., Soza,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,P.,
Williams,G., Willison,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 239076)
Worley,K.C.
Direct Submission
Submitted (14-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 239076)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:22857070.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: G0P1
Center clone name: CH230-87110
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 228642 bases at least Q40
Consensus quality: 232269 bases at least Q30
Consensus quality: 234041 bases at least Q20
Estimated insert size: 231522; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 234710: contig of 234710 bp in length
 * 234711 234810: gap of unknown length
 * 234811 235924: contig of 1114 bp in length
 * 235925 236024: gap of unknown length
 * 236025 237314: contig of 1290 bp in length
 * 237315 237414: gap of unknown length
 * 237415 239076: contig of 1662 bp in length.

FEATURES

Location/Qualifiers
 1..239076
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-87110"
 234711..234810
 /estimated_length=unknown
 235925..236024
 /estimated_length=unknown
 237315..237414
 /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 3.46e+04 Length: 239076
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-8 (1-9) x AC106962 (1-239076)

Qy 1 AlallePheLeuValLeuTyrLeu 9

Db 15610 GCATTTTCCTCCTCGTTTGATTTG 15584

RESULT 45

UBA518259/c UBA518259 387 bp DNA linear ENV 05-DEC-2003

LOCUS Unidentified bacterium partial 16S rRNA gene, clone Mnl292-78.

ACCESSION AJ518259

VERSION AJ518259.1 GI:25265257

KEYWORDS ENV; 16S ribosomal RNA; 16S rRNA gene.

SOURCE unidentified bacterium

ORGANISM

Bacteria; environmental samples.

REFERENCE 1

Wobus, A., Bleul, C., Maassen, S., Scheerer, C., Schuppler, M.,

Jacobs, E. and Rospke, I.

Microbial diversity and functional characterization of sediments

from reservoirs of different trophic state

FEMS Microbiol. Ecol. 46 (3), 331-347 (2003)

REFERENCE 2 (bases 1 to 387)

Bleul, C.

Direct Submission

Submitted (18-NOV-2002) Bleul C., Institute for Medical

Microbiology, Dresden University of Technology, Fiedlerstrasse 42,

Dresden, 01307, GERMANY

Location/Qualifiers

1..387

/organism="unidentified bacterium"

/mol_type="genomic DNA"

/isolation_source="sediment"

/db_xref="taxon:2338"

/clone="Mnl292-78"

/environmental_sample

/country="Germany:Saxony"

1..387

/gene="16S rRNA"

<1..>387

gene

rRNA

/gene="16S rRNA"
 /product="16S ribosomal RNA"

ORIGIN

Alignment Scores:
 Pred. No.: 257 Length: 387
 Score: 40.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 97.6% Indels: 0
 DB: 3 Gaps: 0

US-10-774-176-8 (1-9) x UBA518259 (1-387)

Qy 1 AlallePheLeuValLeuTyrLeu 9

Db 270 GCCGTCTTCCTCTTGACTACTACCTA 244

RESULT 46

AE001270/c

LOCUS

DEFINITION

AE001270

ACCESSION

AE001270

VERSION

AE001270.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Complete genome sequence of *Treponema pallidum*, the syphilis

spirochete

Science 281 (5375), 375-388 (1998)

9665876

REFERENCE 2 (bases 1 to 12448)

Fraser, C.M., Norris, S.J., Weinstein, G.M., White, O., Sutton, G.G.,

Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A.,

Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J.,

Khaklax, H., Richardson, D., Howell, J.K., Chidambaram, M.,

Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M.D.,

Fujii, C., Garland, S., Hatch, B., Hurst, K., Roberts, K., Sandusky, M.,

Weidman, J., Smith, H.O. and Venter, J.C.

Complete genome sequence of *Treponema pallidum*, the syphilis

spirochete

Science 281 (5375), 375-388 (1998)

9665876

REFERENCE 2 (bases 1 to 12448)

Fraser, C.M., Norris, S.J., Weinstein, G.M., White, O., Sutton, G.G.,

Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A.,

Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J.,

Khaklax, H., Richardson, D., Howell, J.K., Chidambaram, M.,

Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M.D.,

Fujii, C., Garland, S., Hatch, B., Hurst, K., Roberts, K., Sandusky, M.,

Weidman, J., Smith, H.O. and Venter, J.C.

Complete genome sequence of *Treponema pallidum*, the syphilis

spirochete

Science 281 (5375), 375-388 (1998)

9665876

REFERENCE 2 (bases 1 to 12448)

Fraser, C.M., Norris, S.J., Weinstein, G.M., White, O., Sutton, G.G.,

Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A.,

Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J.,

Khaklax, H., Richardson, D., Howell, J.K., Chidambaram, M.,

Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M.D.,

Fujii, C., Garland, S., Hatch, B., Hurst, K., Roberts, K., Sandusky, M.,

Weidman, J., Smith, H.O. and Venter, J.C.

Complete genome sequence of *Treponema pallidum*, the syphilis

spirochete

Science 281 (5375), 375-388 (1998)

9665876

REFERENCE 2 (bases 1 to 12448)

Fraser, C.M., Norris, S.J., Weinstein, G.M., White, O., Sutton, G.G.,

Dodson, R., Gwinn, M., Hickey, E.K., Clayton, R., Ketchum, K.A.,

Sodergren, E., Hardham, J.M., McLeod, M.P., Salzberg, S., Peterson, J.,

Khaklax, H., Richardson, D., Howell, J.K., Chidambaram, M.,

Utterback, T., McDonald, L., Artiach, P., Bowman, C., Cotton, M.D.,

Fujii, C., Garland, S., Hatch, B., Hurst, K., Roberts, K., Sandusky, M.,

Weidman, J., Smith, H.O. and Venter, J.C.

Complete genome sequence of *Treponema pallidum*, the syphilis

spirochete

VRVAKXSSPVQAPPLCAVRAQDAQRISCGIAEFDRLVGGAVRRSAIMTGGEPGIG
KSTLLQIAAACSVLVSGESPGQIRGRADRLNIPIONIELLCATRVADHVERVLN
TRCFVTIQTIVSIOQVFSPEACAIPTWLNQKVCANELLAWKURDSVLPTFAHTVKOG
NIAGPKVHEHVDVTIISPERNEEDIRFIRALKRPGSVDELGIPTMGENGISAVQDTA
GFFISTQGMPPVGSATVPVCEGSKRVFVIEQALVPAKSVIRVFSRDIRISVRSRV
AAVTEKRVGLKFSQDDIVYVAGSIRLYEPADVDVALAMALYSARQNTVPKTNAAPIGE
VLSGGEIRPVRLRLKRLNTAYGLQGSTIYVIGVEHDTTPPYTLRVVGTLAQTAEIF
SKAVA"

/complement (1493. .1987)
/gene="TP1023"
/complement (1493. .1987)
/gene="TP1023"
/note="similar to SP:P37861 GB:U05680 PID:468036 percent
identity: 28.89; identified by sequence similarity;
putative"

/codon_start=1
/transl_table=11
/product="recX protein (recX)"
/protein_id="AAC65974.1"
/db_xref="GI:3323352"
/translation="MGTRPTDEQYGVACFACRYEACVAVRLLARSETSQQGLGPKL
LKRGFKRKVSFVLKRYBNLDDTRFARAWLRNRVDSRPASRAQLLGLGKRGVSH
ADAEGLDLLFQEQDEETLCIRLLEKLRGGYGPHTLQALQRQFSPSLVRRCLAVE
TEGA"

/complement (2115. .2504)
/gene="TP1024"
/complement (2115. .2504)
/gene="TP1024"
/note="similar to SP:P07842 percent identity: 54.47;
identified by sequence similarity; putative"

/codon_start=1
/transl_table=11
/product="ribosomal protein S9 (rpsI)"
/protein_id="AAC65981.1"
/db_xref="GI:3323359"
/translation="MNKIGTGRKKTAVARVICRMGMGNVTVNRDDVGVPPTAEOL
RVVREPLPATANERYDIVVNVYGGLDQAGACNAGIRALVRADASNQASLRAGGL
LTRDRMVRKRYQGRGARRPQFSKR"

/complement (2520. .2948)
/gene="TP1025"
/complement (2520. .2948)
/gene="TP1025"
/note="similar to GB:AB000783 percent identity: 54.35;
identified by sequence similarity; putative"

/codon_start=1
/transl_table=11
/product="ribosomal protein L13 (rplM)"
/protein_id="AAC65975.1"
/db_xref="GI:3323353"
/translation="MRTIFVNEREAVRAWHLIDAGRPLGRVAARVACILLRKKHASY
TPNQMGDYVVVINAERKVLSTGTPKDKRKYRHSGYPOGLSGVFSALVGRRPVEPLR
HAVGMPLPKPLGRKLKNVKIYAGSVHPHESQNPVPLSC"

3381. .5150
/gene="TP1026"
3381. .5150
/gene="TP1026"
/note="similar to GB:Z32522 PID:510140 SP:P54124
PID:1771208 percent identity: 24.77; identified by
sequence similarity; putative"

/codon_start=1
/transl_table=11
/product="oligodeopeptidase F, putative"
/protein_id="AAC65976.1"
/db_xref="GI:3323354"

/translation="MQRPEDIPYTPHMDILEHVAADAHGEFAQACARDADAVATILAI
EDLQRLYALANMRCSDIRNTFYRREQDFDAVPRFARLDHAFNQLLASPQR
DGLKLGITHTFARLQSCFCSIMEDLAERNLTISAYETLLASHLFRGHYTL
AQLSPFMHEHTRDAREAYTFPAQHSEBELDLYTLRVLRTRTARTILGYDNYIQ
LGDRLLRSYDMDIARYRTVILRYAVPLAAELHEQQRSLRGLSEILFYDEPLYPFS
GNVPQGDPAWILQAACMYRELSPETQFFTMEYRHLFDVCARIAKASGCYCTLS
TYRAPFIFANFRTAHDVEMTHEVGHAFQAYQRYARLDCLEAYVWPTVEACRIPS
MSMEPLTWMLGFPGEQKRFYLRHLTAQVELLPYGAADVDFQHWVYAHADASATER
KCAWRALESTQPLPRRRYGGHYLSCGGLMMRQSHIFCPFPFYIDYTLAQICALQWDR

gene
CDS

SRVAYTHLSTLTGAAPYASITPTAYARAWHDYCVLCRSGSEPFMRLLATANLHNPFE
EDTFVSTILASCRAYFRITVGDRLS"
5272. .6048
/gene="TP1027"
5272. .6048
/gene="TP1027"
/note="similar to GB:M87049 SP:P12758 GB:X15689 PID:148229
PID:43242 percent identity: 39.18; identified by sequence
similarity; putative"

/codon_start=1
/transl_table=11
/product="uridine phosphorylase (udp)"
/protein_id="AAC65977.1"
/db_xref="GI:3323355"

/translation="MKLVYSTDCEVHIGLKASDIGHVILPGDGPSEKIAQHPFSPH
KVGNREYVYTGILCEPFSVMSTGIGPSTAGVIELHLGAHTFIRVGTSGGQWQ
DILGTVVIATGAIKFEKSTKSEYAPFPAVDFVTALAGAADVDQVRHALGVVQC
KNFYQHSPTMPFVABLTQKHAWIACTNLASEMSAALFVLGSRVRRTGAVLLV
ICNQTTRAQGLIEDIQVHDTENAIRVAVBAVKLLITQDSPR"

/complement (6064. .8076)
/gene="TP1028"
/complement (6064. .8076)
/gene="TP1028"
/note="similar to GP:2618592 percent identity: 40.96;
identified by sequence similarity; putative"

/codon_start=1
/transl_table=11
/product="DNA helicase II (uvrD)"
/protein_id="AAC65978.1"
/db_xref="GI:3323356"

/translation="MESVLSALNEAQRAOVCHYGPSPLILAGAGSGKTRVITTKIAHL
IRSRQVRPQILAVFTWKAAREMTACALESAAGATICTFHALGVWILRYAVRL
GLNPHFSIYDDHDVRAILPKILPHCDHSRAGMLARGISQAKDYGLDCASFESVHARVS
APACAARAVLDQDPFAHAYACVHRHREMGTVDFGLMLPVLLKREHQDVAQLHAR
WQVWVDYQDSNQAQPHFLQVLTGAHTYLCVGGDDQSIYFRGAEVKNILTFPEFF
AALCQILRYENRSTDAIRVADSVKKNQRLKALTAQRTGKTPRLFLNNODEE
AAVQLIQEARAGIPYADWAILRVNAQSLSPFCQFLNRI PYRI VGTLPKFSRAB
VKDVLAFQLIVNGSDELALRRVLNKPFGIGETQDALFVCAQQAIDTFTLOSTH
LTALGTRARQKVSFSLRLRALRMAPQAPAGEARTSAPPEVEETHDAEGLAR
FVSVVMEHTGLEEYRQDEEGTCANVQELMNAASLYACSHGLVSFLHIIQLDQ
NMADEGGAADAVHLITIHTKLEFRVILTGLENGVFPDDEADIEQERRLMYVACT
RAMDSLVTACAYRRMMGRHTAMKSRFLTELDLSALLEITDPRHEA"

/complement (8159. .8845)
/gene="TP1029"
/complement (8159. .8845)
/gene="TP1029"
/note="hypothetical protein; identified by Glimmer;
putative"

/codon_start=1
/transl_table=11
/product="T. pallidum predicted coding region TP1029"
/protein_id="AAC65982.1"
/db_xref="GI:3323360"

/translation="MSKISSKISSVDQSKFOAFLOQILRNVTQEDPRVLDAYRRLFR
KSPFMSRYVAHLAHTHCRAGATAGTRGTGREGITGSAHAPSCASAPASARA
RDFERKARDYPALCPGDTTISIFISGKNRHIYPRDIIALLMORADVAREHIGTIRLD
HYSFIQVLSGEAREAVIARLNGLFYRGRTLTVSHSRRADEHPAPSTEPHAAVAPEPDF
NAEPIPALSE"

/complement (8864. .9361)
/gene="TP1030"
/complement (8864. .9361)
/gene="TP1030"
/note="hypothetical protein; identified by Glimmer;
putative"

/codon_start=1
/transl_table=11
/product="T. pallidum predicted coding region TP1030"
/protein_id="AAC65983.1"
/db_xref="GI:3323361"

/translation="MRGLGHEGERDFOIGGAQPVRLVRLHRRVLFKLNAPLQGV
SYCGAMPSCSEKTSAPNPGSSAPAKTARCSLPTHTFTTTPPPPPVSTLQKQ
PTGACVTAAPAPSSGQGGGALYLFQSVRSGSVSPRAPAAQAQTCPTSHRIM
SARNV"

```

gene          9389..10933

Alignment Scores:
Pred. No.:    4.51e+03      Length:      12448
Score:        40.00         Matches:      8
Percent Similarity: 100.0%   Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match:   97.6%        Indels:      0
DB:            1           Gaps:      0

US-10-774-176-8 (1-9) x AB001270 (1-12448)

Qy 1 Ala1lePheLeuValLeuTyrLeu 9
    |||::|||
Db 3054 GCAGTATTTCTGCTGTTTATATCTC 3028

RESULT 47
AY596297_13
WPCOMMENT
Sequence split into 32 fragments LOCUS AY596297 Accession AY596297
Fragment Name      Begin      End
AY596297_00        1      110000
AY596297_01        100001   210000
AY596297_02        200001   310000
AY596297_03        300001   410000
AY596297_04        400001   510000
AY596297_05        500001   610000
AY596297_06        600001   710000
AY596297_07        700001   810000
AY596297_08        800001   910000
AY596297_09        900001  1010000
AY596297_10       1000001  1110000
AY596297_11       1100001  1210000
AY596297_12       1200001  1310000
AY596297_13       1300001  1410000
AY596297_14       1400001  1510000
AY596297_15       1500001  1610000
AY596297_16       1600001  1710000
AY596297_17       1700001  1810000
AY596297_18       1800001  1910000
AY596297_19       1900001  2010000
AY596297_20       2000001  2110000
AY596297_21       2100001  2210000
AY596297_22       2200001  2310000
AY596297_23       2300001  2410000
AY596297_24       2400001  2510000
AY596297_25       2500001  2610000
AY596297_26       2600001  2710000
AY596297_27       2700001  2810000
AY596297_28       2800001  2910000
AY596297_29       2900001  3010000
AY596297_30       3000001  3110000
AY596297_31       3100001  3131724
Continuation (14 of 32) of AY596297 from base 1300001 (AY596297 Haloarcula marismortui A

Alignment Scores:
Pred. No.:    2.72e+04      Length:      110000
Score:        40.00         Matches:      8
Percent Similarity: 100.0%   Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match:   97.6%        Indels:      0
DB:            1           Gaps:      0

US-10-774-176-8 (1-9) x AY596297_13 (1-110000)

Qy 1 Ala1lePheLeuValLeuTyrLeu 9
    |||::|||
Db 24818 GCAGTATTTCTGCTGTTTATATCTC 24844

RESULT 48
AC121542/c
LOCUS
DEFINITION Mus musculus chromosome 15, clone RP24-137J5, complete sequence.

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

AC121542
AC121542.8 GI:54606954
HTG
Mus musculus (house mouse)
Mus musculus
Mammalia; Euthera; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 164489)
Birren,B., Nusbaum,C. and Lander,B.
Mus musculus chromosome 15, clone RP24-137J5
Unpublished
2 (bases 1 to 164489)
Birren,B., Linton,L., Nusbaum,C., Lander,B., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicoli,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trifilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164489)
Birren,B., Nusbaum,C., Lander,B., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-SEP-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 164489)
Birren,B., Nusbaum,C., Lander,B., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,

```

Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Mlhova, T., Mlenga, V., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, P., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (23-OCT-2004) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Oct 23, 2004 this sequence version replaced gi:52353855.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L24396

Center clone name: L37_J_5

FEATURES

source

1..164489
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="15"
 /map="15"
 /clone="RP24-137J5"
 /clone.lib="RPC1-24 Male Mouse BAC"

repeat_region
 /rpt_family="L1"
 complement(295..677)
 /rpt_family="L1"
 678..751
 /rpt_family="(TTTC)n"
 complement(752..1481)
 /rpt_family="L1_MM"
 1481..1569
 /rpt_family="L1"
 /rpt_family="L1"
 2554..2590
 /rpt_family="AT_rich"
 2659..2820
 /rpt_family="MTD"
 2825..2899
 /rpt_family="(TCTCTG)n"
 3666..3745
 /rpt_family="MTG"
 3809..3857
 /rpt_family="AT_rich"
 3867..3888
 /rpt_family="(T)n"
 4393..4499
 /rpt_family="RSINE1"
 4512..4586
 /rpt_family="(CA)n"
 complement(5848..5956)
 /rpt_family="Lx"
 6327..6355
 /rpt_family="AT_rich"
 6383..6498
 /rpt_family="PB1D7"
 complement(6824..7694)
 /rpt_family="Lx8"
 complement(7993..8292)
 /rpt_family="LIMA9"

repeat_region
 complement(8429..8775)
 /rpt_family="L1P5"
 repeat_region
 complement(8810..9256)
 /rpt_family="L1P5"
 9257..9301
 /rpt_family="(CA)n"
 9301..9347
 /rpt_family="(TA)n"
 /rpt_family="(CA)n"
 complement(9302..9967)
 /rpt_family="L1P5"
 9968..10131
 /rpt_family="ORR1D"
 10298..10371
 /rpt_family="(CAAAA)n"
 complement(10591..10752)
 /rpt_family="L1_MM"
 10971..10997
 /rpt_family="(TTTG)n"
 12010..12217
 /rpt_family="(CAAA)n"
 /rpt_family="Lx7"
 complement(12540..12944)
 /rpt_family="Lx6"
 12945..13038
 /rpt_family="(TTTG)n"
 complement(13055..13284)
 /rpt_family="Lx6"
 13333..14048
 /rpt_family="RMER1B"
 complement(14264..14356)
 /rpt_family="MIR"
 14835..15021
 /rpt_family="LIMA9"
 complement(15304..15441)
 /rpt_family="RSINE1"
 15456..15491
 /rpt_family="AT_rich"
 complement(15558..15672)
 /rpt_family="S8"
 complement(15917..16040)
 /rpt_family="RLTR9C"
 16043..16401
 /rpt_family="Lx8"
 16614..16670
 /rpt_family="(TG)n"
 16689..16868
 /rpt_family="B2_Mm2"
 17740..17789
 /rpt_family="AT_rich"
 complement(17790..17903)
 /rpt_family="RMER1B"
 17921..17981
 /rpt_family="(TC)n"

Alignment Scores:

Pred. No.: 3.79e+04 Length: 164489
 Score: 40.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 97.6% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-8 (1-9) x AC121542 (1-164489)

QY 1 AlaiPheLeuValleyrleu 9

Db 45724 GCAGTTTCTTTGGTATATCTA 45698
 |||:|||||
 |||:|||||

RESULT 49

AC163932

LOCUS

DEFINITION

pieces.

ACCESSION

AC163932

Bos taurus clone CH240-107F10, WORKING DRAFT SEQUENCE, 9 unordered

linear

HTG 01-JUL-2005

AC163932.2 GI:68300994
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTDP.
Bos taurus
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 212987)
Muzny,D., Marie., Metzker, M., Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P., Bryant, K., Buhay, C., Burch, P., Burrell, K., Blyth, P., Brown, M., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, T., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Loushew, L., Louised, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B., Mawney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankemeh, O., Okwuonu, G., Olarnpungsoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Fu, L., L., Pua, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseas, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Unpublished
Direct Submission
2 (bases 1 to 212987)
Worley, K. C.
Submitted (15-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 212987)
Cow Genome Sequencing Consortium.
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 29, 2005 this sequence version replaced gi:67763874.
The sequence in this assembly is a combination of BAC based reads

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: FHP
Center clone name: CH240-107F10
Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 209389 bases at least Q40
Consensus quality: 210366 bases at least Q30
Consensus quality: 211128 bases at least Q20
Estimated insert size: 212035; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2952: contig of 2952 bp in length
2953 3149: gap of 197 bp
3150 34920: contig of 31771 bp in length
34921 34970: gap of 50 bp
34971 54700: contig of 19730 bp in length
54701 54750: gap of 50 bp
54751 56234: contig of 1484 bp in length
56235 56528: gap of 294 bp
56529 74429: contig of 17901 bp in length
74430 74479: gap of 50 bp
74480 85346: contig of 10867 bp in length
85347 85396: gap of 50 bp
85397 102085: contig of 16689 bp in length
102086 102135: gap of 50 bp
102136 211461: contig of 109326 bp in length
211462 211561: gap of unknown length
211562 212987: contig of 1426 bp in length.

Location/Qualifiers
1. 212987
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-107F10"
2953. 3149
/estimated length=197
34921. 34970
/estimated length=50
54701. 54750
/estimated length=50
56235. 56528
/estimated length=294
74430. 74479
/estimated length=50
85347. 85396
/estimated length=50
102086. 102135

FEATURES
source

gap
gap
gap
gap
gap
gap
gap

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

gap /estimated_length=50
211462..211561
/estimated_length-unknown

ORIGIN

Alignment Scores:

Pred. No.: 4.7e+04 Length: 212987
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 97.6% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-8 (1-9) x AC163932 (1-212987)

Qy 1 AlallePheLeuValLeuTyRieu 9

Db 49509 GCCATTTTCCTTCATTCGTACCTC 49535

RESULT 50

AC101062 AC101062 63024 bp DNA linear HTG 23-NOV-2001
LOCUS Mus musculus clone RP23-87F10, LOW-PASS SEQUENCE SAMPLING.

ACCESSION

AC101062

VERSION

AC101062.1 GI:17059836

KEYWORDS

HTG, HTGS_PHASE0.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 63024)

Biren, B., Linton, L., Nusbaum, C. and Lander, E.

Mus musculus, clone RP23-87F10

Unpublished

2 (bases 1 to 63024)

Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,

Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,

Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,

Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,

Ferreira, P., FitzHugh, W., Gage, D., Galagen, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,

Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,

Jones, C., Katat, A., Karas, A., Kells, C., Larocque, K.,

Lamazares, R., Landers, P., Lehoczy, J., Levine, R., Liu, G.,

MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,

McCarthy, M., McGowan, P., McKernan, K., McPheeters, R., Meldrim, J.,

Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,

Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P.,

Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,

Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIER

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L14879

Center Clone name: 87_F_10

* NOTE: This record contains 74 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1 756: contig of 756 bp in length
* 757 856: gap of 100 bp
* 857 1628: contig of 772 bp in length
* 1629 1728: gap of 100 bp
* 1729 2497: contig of 769 bp in length
* 2498 2597: gap of 100 bp
* 2598 3360: contig of 763 bp in length
* 3361 3460: gap of 100 bp
* 3461 4199: contig of 739 bp in length
* 4200 4299: gap of 100 bp
* 4300 5058: contig of 759 bp in length
* 5059 5158: gap of 100 bp
* 5159 5918: contig of 760 bp in length
* 5919 6018: gap of 100 bp
* 6019 6771: contig of 753 bp in length
* 6772 6871: gap of 100 bp
* 6872 7584: contig of 713 bp in length
* 7585 7684: gap of 100 bp
* 7685 8447: contig of 763 bp in length
* 8448 8547: gap of 100 bp
* 8548 9315: contig of 768 bp in length
* 9316 9415: gap of 100 bp
* 9416 10162: contig of 747 bp in length
* 10163 10262: gap of 100 bp
* 10263 11015: contig of 753 bp in length
* 11016 11115: gap of 100 bp
* 11116 11872: contig of 757 bp in length
* 11873 11972: gap of 100 bp
* 11973 12738: contig of 766 bp in length
* 12739 12838: gap of 100 bp
* 12839 13594: contig of 756 bp in length
* 13595 13694: gap of 100 bp
* 13695 14450: contig of 756 bp in length
* 14451 14550: gap of 100 bp
* 14551 15301: contig of 751 bp in length
* 15302 15401: gap of 100 bp
* 15402 16179: contig of 778 bp in length
* 16180 16279: gap of 100 bp
* 16280 17037: contig of 758 bp in length
* 17038 17137: gap of 100 bp
* 17138 17893: contig of 756 bp in length
* 17894 17993: gap of 100 bp
* 17994 18734: contig of 741 bp in length
* 18735 18834: gap of 100 bp
* 18835 19598: contig of 764 bp in length
* 19599 19698: gap of 100 bp
* 19699 20452: contig of 754 bp in length
* 20453 20552: gap of 100 bp
* 20553 21312: contig of 760 bp in length
* 21313 21412: gap of 100 bp
* 21413 22145: contig of 733 bp in length
* 22146 22245: gap of 100 bp
* 22246 23006: contig of 761 bp in length
* 23007 23106: gap of 100 bp
* 23107 23856: contig of 750 bp in length
* 23857 23956: gap of 100 bp
* 23957 24722: contig of 766 bp in length
* 24723 24822: gap of 100 bp
* 24823 25577: contig of 755 bp in length
* 25578 25677: gap of 100 bp
* 25678 26410: contig of 733 bp in length
* 26411 26510: gap of 100 bp

* 26511 27263: contig of 753 bp in length
* 27264 27363: gap of 100 bp
* 27364 28111: contig of 748 bp in length
* 28111 28211: gap of 100 bp
* 28211 28974: contig of 763 bp in length
* 28974 29075: gap of 100 bp
* 29075 29831: contig of 757 bp in length
* 29831 29931: gap of 100 bp
* 29931 30668: contig of 737 bp in length
* 30668 30769: gap of 100 bp
* 30769 31531: contig of 763 bp in length
* 31531 31631: gap of 100 bp
* 31631 32395: contig of 764 bp in length
* 32395 32495: gap of 100 bp
* 32495 33263: contig of 768 bp in length
* 33263 33363: gap of 100 bp
* 33363 34127: contig of 764 bp in length
* 34127 34227: gap of 100 bp
* 34227 34989: contig of 762 bp in length
* 34989 35089: gap of 100 bp
* 35089 35831: contig of 742 bp in length
* 35831 36689: contig of 758 bp in length
* 36689 36789: gap of 100 bp
* 36789 37535: contig of 746 bp in length
* 37535 37635: gap of 100 bp
* 37635 38379: contig of 744 bp in length
* 38379 38479: gap of 100 bp
* 38479 39220: contig of 741 bp in length
* 39220 39320: gap of 100 bp
* 39320 40058: contig of 738 bp in length
* 40058 40158: gap of 100 bp
* 40158 40927: contig of 769 bp in length
* 40927 41027: gap of 100 bp
* 41027 41746: contig of 719 bp in length
* 41746 41846: gap of 100 bp
* 41846 42587: contig of 741 bp in length
* 42587 42687: gap of 100 bp
* 42687 43436: contig of 749 bp in length
* 43436 43536: gap of 100 bp
* 43536 44267: contig of 731 bp in length
* 44267 44367: gap of 100 bp
* 44367 45129: contig of 762 bp in length
* 45129 45229: gap of 100 bp
* 45229 45971: contig of 742 bp in length
* 45971 46071: gap of 100 bp
* 46071 46819: contig of 748 bp in length
* 46819 46919: gap of 100 bp
* 46919 47699: contig of 780 bp in length
* 47699 47799: gap of 100 bp
* 47799 48557: contig of 758 bp in length
* 48557 48657: gap of 100 bp
* 48657 49400: contig of 743 bp in length
* 49400 49500: gap of 100 bp
* 49500 50250: contig of 750 bp in length
* 50250 51116: contig of 766 bp in length
* 51116 51216: gap of 100 bp
* 51216 51980: contig of 764 bp in length
* 51980 52080: gap of 100 bp
* 52080 52835: contig of 755 bp in length
* 52835 52935: gap of 100 bp
* 52935 53696: contig of 761 bp in length
* 53696 53796: gap of 100 bp
* 53796 54559: contig of 763 bp in length
* 54559 54659: gap of 100 bp
* 54659 55402: contig of 743 bp in length
* 55402 55502: gap of 100 bp
* 55502 56256: contig of 754 bp in length
* 56256 56356: gap of 100 bp
* 56356 57084: contig of 728 bp in length
* 57084 57184: gap of 100 bp
* 57184 57939: contig of 755 bp in length
* 57939

* 57940 58039: gap of 100 bp
* 58040 58778: contig of 739 bp in length

Alignment Scores:
Pred. No.: 2.57e+04 Length: 63024
Score: 39.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.1% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-8 (1-9) x AC101062 (1-63024)

Qy 1 A1a1lePheLeuLeuValLeuTyrLeu 9

Db 59223 GCAATATTTCTCTGGTCATATATCTA 59249

Search completed: April 25, 2006, 20:33:12

Job time : 3094.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds

(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-7

Perfect score: 43

Sequence: 1 SLQTSYVFL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/abss/ABSSWEB spool/US10774176/runat_24042006.165112.19185/app_query.fasta.1
-DB=N Geneseq -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT5-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss05p
-USER=US10774176 @CGN 1.1 3463 @runat_24042006.165112.19185 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -PGAPOP=6 -PGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 21.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	43	100.0	65	6	Abn56274 Mouse spl
2	43	100.0	108	10	ACD97670 Human col
3	43	100.0	453	5	Aas87174 DNA encod
4	43	100.0	475	13	ADU11677 Solid tum

5	43	100.0	901	3	AA27060
6	43	100.0	927	6	ABT07721
7	43	100.0	927	8	ABX76333
8	43	100.0	927	10	ADB80503
9	43	100.0	927	11	ADN38723
10	43	100.0	973	8	AAD56198
11	43	100.0	1156	6	ABV99349
12	43	100.0	1260	6	ABK87175
13	43	100.0	1260	10	ADB97513
14	43	100.0	1260	10	ADB97452
15	43	100.0	1263	3	AAA27058
16	43	100.0	1263	4	AAF89736
17	43	100.0	1263	6	ABK87174
18	43	100.0	1281	3	AA27059
19	43	100.0	1331	8	AA256199
20	43	100.0	2020	10	ADJ56299
21	43	100.0	2053	8	ACC51052
22	43	100.0	2053	8	ABX76332
23	43	100.0	2053	8	AAD56197
24	43	100.0	2053	8	AAD56200
25	43	100.0	2053	11	ADN38721
26	43	100.0	2053	12	ADL06473
27	43	100.0	2053	12	ADN03961
28	43	100.0	2053	13	ADR25444
29	43	100.0	2053	13	ACN38510
30	43	100.0	2053	13	ADV35098
31	43	100.0	2338	5	AAS87175
32	43	100.0	2359	4	AAK94253
33	43	100.0	2359	12	ADL30831
34	43	100.0	2361	4	AAK94254
35	43	100.0	2361	12	ADI26162
36	43	100.0	2361	12	ADL30833
37	43	100.0	2557	12	ADI26160
38	43	100.0	2557	12	ADI26158
39	90.7	14952	10	ACF69283	
40	39	90.7	100779	10	ACF65386
41	39	90.7	110000	10	ACF67367
42	38	88.4	117730	14	ADZ12550
43	37	86.0	565	13	ADX45743
44	37	86.0	2848	2	AAV75704
45	37	86.0	30346	8	AAD56099
46	37	86.0	30346	9	ADA02461
47	37	86.0	30346	10	ADB72200
48	37	86.0	30346	10	ADB82932
49	37	86.0	176594	13	ABD33387
50	36	83.7	199	6	ABL75414
51	36	83.7	270	6	ABL39975
52	36	83.7	506	9	ACH33681
53	36	83.7	512	9	ACH37372
54	36	83.7	824	4	AAI96322
55	36	83.7	837	3	AAZ35834
56	36	83.7	837	3	AA59874
57	36	83.7	898	5	AAS87139
58	36	83.7	1000	14	ABE85710
59	36	83.7	1155	6	AAL39976
60	36	83.7	1564	12	ADP04766
61	36	83.7	2444	6	ABZ11756
62	36	83.7	2444	12	ADM44274
63	36	83.7	2535	12	ADN05055
64	36	83.7	2535	13	ADP55416
65	36	83.7	12932	10	ADC85672
66	36	83.7	23626	4	AAS28506
67	36	83.7	23626	5	ABA21480
68	36	83.7	23626	5	AAS29997
69	36	83.7	23626	10	ADB33334
70	36	83.7	23626	10	ADG41702
71	36	83.7	23626	11	ADI97476
72	36	83.7	23632	4	AAS28507
73	36	83.7	23632	5	ABA21481
74	36	83.7	23632	5	AAS29998
75	36	83.7	23632	10	ADB33335
76	36	83.7	23632	10	ADG41703
77	36	83.7	23632	11	ADI97477

78	36	83.7	110000	6	ABX08336_01	Continuation (2 of	151	34	79.1	3099	4	ABL04822
79	36	83.7	110000	12	ADJ25985_01	Continuation (2 of	c 152	34	79.1	3262	4	ABL19889
80	36	83.7	110000	12	ADN97989_01	Continuation (2 of	153	34	79.1	4285	2	AAV55749
81	36	83.7	110000	12	ADOS0288_01	Continuation (2 of	154	34	79.1	4285	3	ABQ92075
82	36	83.7	110000	14	ABH85185_01	Continuation (2 of	155	34	79.1	4540	3	AAAB1739
83	36	83.4	110	14	ADX04445	Adx04445 Rat prima	156	34	79.1	5109	10	ADP81540
84	35	81.4	110	14	ADX04445	Adx044289 Mouse pri	157	34	79.1	5109	13	ADP81540
85	35	81.4	199	3	AAC16458	Aac16458 Human sec	158	34	79.1	5261	5	ADL45687
86	35	81.4	684	6	ABZ13158	Abz13158 Arabidops	159	34	79.1	5776	4	ABL19888
87	35	81.4	684	6	ADG87922	Adg87922 A. thalia	c 160	34	79.1	5989	4	AAU46535
88	35	81.4	684	8	ADA68387	Ada68387 Arabidops	161	34	79.1	6043	4	ABL03598
89	35	81.4	849	3	AAC38433	Aac38433 Arabidops	c 162	34	79.1	7504	4	AAK76561
90	35	81.4	995	3	AAC47255	Aac47255 Arabidops	c 163	34	79.1	10538	4	AAK71513
91	35	81.4	1281	8	ACN44984	Acn44984 Prokaryot	164	34	79.1	15831	4	ABL10134
92	35	81.4	1293	10	ADP03273	Adp03273 Bacterial	165	34	79.1	18998	4	AAU36452
c 93	35	81.4	1410	3	AAC41777	Aac41777 Arabidops	166	34	79.1	18998	10	ADL47146
c 94	35	81.4	2000	8	ADA72406	Ada72406 Rice gene	167	34	79.1	18998	13	ADJ08564
c 95	35	81.4	2000	11	ACL35594	ACL35594 Rice stre	c 168	34	79.1	19459	6	ABX31212
c 96	35	81.4	2000	12	ADJ41598	Adj41598 Plant cDN	c 169	34	79.1	19459	6	ABL70527
c 97	35	81.4	6223	6	ASG61176	Asg61176 Human gen	170	34	79.1	26371	11	ACN44814
c 98	35	81.4	6412	13	ABD33380	Abd33380 Murine ca	171	34	79.1	32191	4	AAU36455
c 99	35	81.4	7696	4	AAU46399	AAU46399 Tumour su	172	34	79.1	32191	10	ADL47149
c 100	35	81.4	10006	6	ABL32036	Abi32036 Human imm	173	34	79.1	32191	13	ADJ08567
c 101	35	81.4	32217	4	AAU41738	AAU41738 Genomic s	c 174	34	79.1	33393	13	ABD33223
c 102	35	81.4	61396	13	ABD33379	Abd33379 Murine ca	c 175	34	79.1	42061	13	ABD33566
c 103	35	81.4	65047	11	ACN44020	Acn44020 Mouse gen	176	34	79.1	46456	9	ADA02825
c 104	35	81.4	89328	6	ABL61995	Abi61995 Colon ade	177	34	79.1	46456	10	ADL72563
c 105	35	81.4	98642	11	ACN44584	Adp08391 Human tit	178	34	79.1	46456	12	ADM74420
c 106	35	81.4	98642	11	ACN44584	Adp08391 Human tit	179	34	79.1	46456	10	ADL74420
c 107	35	81.4	110000	6	ABA90193_2	Continuation (3 of	c 180	34	79.1	52754	9	ADA02798
c 108	35	81.4	110000	6	ABQ87681_2	Continuation (3 of	c 181	34	79.1	52754	10	ADL72536
c 109	35	81.4	110000	8	ABX33717_2	Continuation (3 of	c 182	34	79.1	52754	10	ADL72536
c 110	35	81.4	110000	10	ADH81391_1	Continuation (2 of	c 183	34	79.1	52754	12	ADM85278
c 111	35	81.4	110000	11	ACN44932_0	Continuation (8 of	c 184	34	79.1	55829	13	ABD33512
c 112	35	81.4	202351	11	ACN44504_07	Continuation (8 of	c 185	34	79.1	55829	14	ADL12653
c 113	35	81.4	202351	11	ACN44504_07	Continuation (8 of	c 186	34	79.1	76138	12	ADQ97334
c 114	35	81.4	330973	11	ACN44846	Acn44846 Human gen	c 187	34	79.1	80963	14	AEA61095_3
c 115	34	79.1	118	4	AAU00100	AAU00100 Human rep	188	34	79.1	110000	3	AAU81489_1
c 116	34	79.1	118	4	ABA07466	ABA07466 Human ova	189	34	79.1	110000	12	ADQ97047_1
c 117	34	79.1	159	10	ABZ38318	Abz38318 N. gonorr	190	34	79.1	110000	12	ADQ97138_2
c 118	34	79.1	400	6	ABN20456	Abn20456 Human ORF	191	34	79.1	110000	12	ADQ97138_3
c 119	34	79.1	441	9	ACH24713	Ach24713 Human adu	192	34	79.1	113240	11	ACN45082
c 120	34	79.1	470	5	ABV58437	Abv58437 Human pro	193	34	79.1	116219	14	ADL13297
c 121	34	79.1	490	3	AAC51822	Aac51822 Arabidops	c 194	34	79.1	171398	14	ADL13297
c 122	34	79.1	492	10	ADH81849	Adh81849 Arabidops	195	34	79.1	175390	10	ADL13297
c 123	34	79.1	504	10	ADP79592	Adp79592 Leukaemia	196	34	79.1	175390	14	ADQ50650
c 124	34	79.1	524	4	AAU02953	AAU02953 Human rep	c 197	34	79.1	185548	13	ADV77908
c 125	34	79.1	524	4	ABA07727	Abv07727 Human ova	c 198	34	79.1	185548	13	ABD33586
c 126	34	79.1	544	5	ABV52539	Abv52539 Human pro	c 199	34	79.1	191584	13	ADR67026
c 127	34	79.1	565	13	ACN48041	Acn48041 Cotton pr	200	34	79.1	191584	13	ADR67026
c 128	34	79.1	634	6	ABK36022	Abk36022 cDNA sequ	201	34	79.1	209612	12	ADQ59395
c 129	34	79.1	654	12	ADJ48179	Adj48179 Maize oil	c 202	34	79.1	209612	14	ADL13662
c 130	34	79.1	666	13	ADT45645	Adt45645 Bacterial	203	34	79.1	215974	12	ADQ97523
c 131	34	79.1	670	3	ACN434329	Acn434329 Arabidops	c 204	34	79.1	330973	11	ACN44846
c 132	34	79.1	671	3	AAC37355	Aac37355 Arabidops	c 205	33	76.7	349980	3	AAU21611
c 133	34	79.1	672	3	ACN35458	Aac35458 Arabidops	c 206	33	76.7	65	6	ABN28372
c 134	34	79.1	703	3	AAC40739	Aac40739 Arabidops	c 207	33	76.7	100	8	ACD75155
c 135	34	79.1	786	13	ADH46589	Adh46589 Bacterial	c 208	33	76.7	187	3	AAU00954
c 136	34	79.1	845	12	ADQ61171	Adq61171 Tomato ke	c 209	33	76.7	298	10	ABX82490
c 137	34	79.1	889	4	AAU01310	AAU01310 Human rep	210	33	76.7	331	6	ABQ97643
c 138	34	79.1	1212	13	ADX60026	Adx60026 Plant ful	c 211	33	76.7	369	6	ABQ97643
c 139	34	79.1	1391	13	ADX52438	Adx52438 Plant ful	c 212	33	76.7	397	4	AAI84668
c 140	34	79.1	1394	4	AAK73651	AAK73651 Human imm	c 213	33	76.7	443	6	ABL62132
c 141	34	79.1	1514	3	AAK95492	Aac95492 Human sec	214	33	76.7	494	4	AAI81409
c 142	34	79.1	1569	13	ADX59709	Adx59709 Plant ful	c 215	33	76.7	500	4	AAU05670
c 143	34	79.1	2000	8	ADA69310	Ada69310 Arabidops	c 216	33	76.7	500	4	AAU05668
c 144	34	79.1	2000	8	ADA72619	Ada72619 Rice gene	c 217	33	76.7	500	9	ACH47074
c 145	34	79.1	2211	10	ACL37109	ACL37109 Rice stre	c 218	33	76.7	510	10	ADL32271
c 146	34	79.1	2211	4	ABA09064	ABA09064 Human sec	c 219	33	76.7	510	12	ACH74260
c 147	34	79.1	2242	4	ABL23778	Abi23778 Drosophil	c 220	33	76.7	523	5	ABV29725
c 148	34	79.1	2242	13	ADR08381	Adr08381 Full leng	c 221	33	76.7	523	5	ABV23848
c 149	34	79.1	2361	6	AAU40743	AAU40743 Human kin	c 222	33	76.7	527	5	ABV15197
c 150	34	79.1	2477	11	ADN02480	Adn02480 Thiamine	c 223	33	76.7	530	5	ABV06028

C 370	33	76.7	71979	13	ADV87736	Adv87736 Streptococ	443	74.4	564	11	ACL29689	ACL29689 Rice abio
C 371	33	76.7	71979	13	ADV79898	Adv79898 Streptococ	444	74.4	580	5	AAA14043	AAA14043 Human ner
C 372	33	76.7	77001	14	ADZ42277	Adz42277 Human hep	C 445	74.4	594	3	AA667184	AA667184 Pinus rad
C 373	33	76.7	88892	12	ADQ97695	Adq97695 Human can	C 446	74.4	595	6	ABN84546	ABN84546 Human col
C 374	33	76.7	89060	14	AEA61142	Aea61142 Human ABC	C 447	74.4	595	12	ADO41057	ADO41057 Human col
C 375	33	76.7	92562	10	ADC85284	Adc85284 Human itk	C 448	74.4	598	6	ABN84545	ABN84545 Human col
C 376	33	76.7	92563	9	ADA02804	Ada02804 Human itk	C 449	74.4	658	4	AA846890	AA846890 Human G p
C 377	33	76.7	92563	10	ADB72542	Adb72542 Human itk	C 450	74.4	658	4	ABK81668	ABK81668 cDNA enco
C 378	33	76.7	92563	12	ADM74399	Adm74399 Human car	C 451	74.4	660	6	ABQ67641	ABQ67641 Listeria
C 379	33	76.7	96599	9	ADA02981	Ada02981 Mouse Map	C 452	74.4	663	6	ABC32773	ABC32773 Arabidops
C 380	33	76.7	96599	10	ADB72719	Adb72719 Mouse Map	C 453	74.4	673	6	ABK63243	ABK63243 Rat sequ
C 381	33	76.7	96599	10	ADC85461	Adc85461 Mouse Map	C 454	74.4	708	6	ABQ69708	ABQ69708 Listeria
C 382	33	76.7	96599	12	ADM74576	Adm74576 Murine ca	C 455	74.4	710	3	AAC49417	AAC49417 Arabidops
C 383	33	76.7	96861	12	ADQ97834	Adq97834 Human can	C 456	74.4	712	3	ABC34715	ABC34715 Arabidops
C 384	33	76.7	110000	6	ABN71527	Continuation (9 of	C 457	74.4	713	6	ABE52774	ABE52774 Murine tu
C 385	33	76.7	110000	6	ABN71527	Continuation (2 of	C 458	74.4	714	10	ABX07256	ABX07256 S. pneumo
C 386	33	76.7	110000	11	ACN44932	Continuation (2 of	C 459	74.4	714	10	AA555894	AA555894 Streptoco
C 387	33	76.7	110000	13	ABD32909	Continuation (6 of	C 460	74.4	717	4	AA555773	AA555773 Streptoco
C 388	33	76.7	110000	13	ABD32909	Continuation (5 of	C 461	74.4	717	8	ABZ42276	ABZ42276 Streptoco
C 389	33	76.7	110000	13	ADV81204	Continuation (10 of	C 462	74.4	717	8	ACA50017	ACA50017 Prokaryot
C 390	33	76.7	110000	13	ADZ12814	Continuation (3 of	C 463	74.4	717	13	ADR93069	ADR93069 Novel S.
C 391	33	76.7	110000	14	ABE43401	Continuation (4 of	C 464	74.4	717	13	ADK44205	ADK44205 Streptoco
C 392	33	76.7	110096	6	ABN95044	Abn95044 Gene #154	C 465	74.4	717	13	AEA56939	AEA56939 Streptoco
C 393	33	76.7	119950	2	AXX90201	Adx90201 Human yes	C 466	74.4	754	6	ABL34309	ABL34309 Human imm
C 394	33	76.7	123192	13	ADV34995	Adv34995 Murine cd	C 467	74.4	768	6	ABN68653	ABN68653 Streptoco
C 395	33	76.7	128668	11	ACN44074	Acn44074 Human gen	C 468	74.4	770	9	AA160854	AA160854 Human CYP
C 396	33	76.7	133894	2	AA113635	AA113635 ACNPV gen	C 469	74.4	771	8	ACA50648	ACA50648 Prokaryot
C 397	33	76.7	144524	14	AEA62867	Aea62867 Rat glyco	C 470	74.4	771	14	ADZ38605	ADZ38605 Group A S
C 398	33	76.7	146547	8	ABZ80817	Abz80817 Human pho	C 471	74.4	831	13	ADS54460	ADS54460 Bacteri
C 399	33	76.7	155350	13	ABD33514	Abd33514 Murine ca	C 472	74.4	832	4	AA196112	AA196112 Human neu
C 400	33	76.7	159400	6	ABQ88126	Abq88126 Human oat	C 473	74.4	861	5	AA568906	AA568906 DNA encod
C 401	33	76.7	170245	10	ADP13586	Adp13586 Renal cel	C 474	74.4	868	8	AA160073	AA160073 Heteracti
C 402	33	76.7	186510	12	ADP24797	Adp24797 Human end	C 475	74.4	877	13	ADT16180	ADT16180 Plant cDN
C 403	33	76.7	205388	12	ADQ97560	Adq97560 Mouse can	C 476	74.4	904	3	AAF88266	AAF88266 H. tuberc
C 404	33	76.7	210920	12	ADQ97123	Adq97123 Mouse can	C 477	74.4	912	3	AAH31095	AAH31095 Human col
C 405	33	76.7	235033	2	AAV57926	AAv57926 Hereditar	C 478	74.4	987	10	ADC26788	ADC26788 Human lip
C 406	33	76.7	237326	2	AAV57903	AAv57903 Hereditar	C 479	74.4	993	6	ABQ68430	ABQ68430 Listeria
C 407	33	76.7	254481	12	ADQ97135	Adq97135 Mouse can	C 480	74.4	1017	10	ABE09559	ABE09559 Novel DNA
C 408	33	76.7	263744	10	ADF08271	Adf08271 Mouse apo	C 481	74.4	1074	2	AA160095	AA160095 Serpin pr
C 409	33	74.4	229	3	AA160259	AA160259 Human sec	C 482	74.4	1085	3	AA44815	AA44815 Arabidops
C 410	33	74.4	255	5	AAH82030	AAh82030 Rat diff	C 483	74.4	1146	8	ACA37829	ACA37829 Prokaryot
C 411	33	74.4	255	6	ABN92966	Abn92966 Staphyloc	C 484	74.4	1188	13	ADS45544	ADS45544 Bacteri
C 412	33	74.4	255	13	AD802177	Ad802177 Staphyloc	C 485	74.4	1200	1	AAH90713	AAH90713 Gene I of
C 413	33	74.4	281	3	AAH43792	AAh43792 Human sec	C 486	74.4	1218	6	ABN92959	ABN92959 Staphyloc
C 414	33	74.4	300	10	ADC26824	Adc26824 Human lip	C 487	74.4	1218	13	AD802163	AD802163 Staphyloc
C 415	33	74.4	319	4	AAK55038	AAk55038 Human imm	C 488	74.4	1259	6	ABN95869	ABN95869 Human dih
C 416	33	74.4	322	6	ABL78023	ABl78023 Human ova	C 489	74.4	1261	9	ACC84973	ACC84973 CYP3A4E7-
C 417	33	74.4	331	4	AAH34200	AAh34200 Human col	C 490	74.4	1284	10	ADK61939	ADK61939 Disease t
C 418	33	74.4	343	4	AAK80578	AAk80578 Human imm	C 491	74.4	1294	6	ABZ16449	ABZ16449 Arabidops
C 419	33	74.4	343	4	AAK80579	AAk80579 Human imm	C 492	74.4	1294	8	ACA34739	ACA34739 Prokaryot
C 420	33	74.4	374	4	AA184793	AA184793 Human pol	C 493	74.4	1333	10	ADI22454	ADI22454 Rat liver
C 421	33	74.4	378	8	ABX43040	ABx43040 Bovine ES	C 494	74.4	1338	2	AAH14340	AAH14340 H. pylori
C 422	33	74.4	384	5	AA666885	AAf666885 Novel hum	C 495	74.4	1347	13	ADS46299	ADS46299 Bacteri
C 423	33	74.4	411	4	AA191406	AA191406 Human pol	C 496	74.4	1383	13	ADT43518	ADT43518 Bacteri
C 424	33	74.4	424	5	ABV15934	ABv15934 Human pro	C 497	74.4	1393	5	AA529901	AA529901 Human lun
C 425	33	74.4	429	2	AAH28300	AAh28300 Human CYP	C 498	74.4	1393	10	ADW33146	ADW33146 Human nov
C 426	33	74.4	430	6	ABL84931	ABl84931 Human ova	C 499	74.4	1406	14	ABM16483	ABM16483 Eucalyptu
C 427	33	74.4	446	5	ABV45734	ABv45734 Human pro	C 500	74.4	1491	3	AACT76525	AACT76525 Human ORF
C 428	33	74.4	447	13	ACF91498	ACf91498 Human SIR	C 501	74.4	1505	13	ADX63461	ADX63461 Plant ful
C 429	33	74.4	453	3	AAH30687	AAh30687 Human col	C 502	74.4	1542	13	ADX63785	ADX63785 Arabidops
C 430	33	74.4	457	9	AA186357	AA186357 Human pol	C 503	74.4	1660	3	AAQ75336	AAQ75336 Superoxid
C 431	33	74.4	462	9	ACH43793	ACH43793 Human foe	C 504	74.4	1729	2	AAQ75336	AAQ75336 Superoxid
C 432	33	74.4	467	13	ADT17770	Adt17770 Plant cDN	C 505	74.4	1729	6	ABK63674	ABK63674 Rat sequ
C 433	33	74.4	504	6	ABK62869	ABk62869 Rat sequ	C 506	74.4	1729	10	AD858137	AD858137 Toxicity-
C 434	33	74.4	504	10	ADB56866	ADB56866 Toxicity-	C 507	74.4	1729	10	ABT41857	ABT41857 Toxicity
C 435	33	74.4	504	10	ABT41233	ABt41233 Toxicity-	C 508	74.4	1729	12	ADP73021	ADP73021 Renal tox
C 436	33	74.4	504	13	ADV40283	Adv40283 Rat cardi	C 509	74.4	1821	5	ADL62684	ADL62684 Human ova
C 437	33	74.4	510	9	ADA30864	Ada30864 DNA encod	C 510	74.4	1827	10	ADK60171	ADK60171 Plant DNA
C 438	33	74.4	525	3	AA94728	AAa94728 Soybean a	C 511	74.4	1893	8	ACA48435	ACA48435 Prokaryot
C 439	33	74.4	540	4	AAK61900	AAk61900 Human imm	C 512	74.4	1965	6	ABQ70493	ABQ70493 Listeria
C 440	33	74.4	542	6	ABT11076	ABt11076 Human bre	C 513	74.4	2000	6	ABZ17131	ABZ17131 Arabidops
C 441	33	74.4	551	12	ADL84276	ADl84276 DNA up-re	C 514	74.4	2000	8	ADA71696	ADA71696 Rice gene
C 442	33	74.4	551	12	ADL84275	ADl84275 DNA up-re	C 515	74.4	2000	11	ACL36481	ACL36481 Rice stre

C 516	32	74.4	2000	11	ACL37402	AcL37402 Rice stre	589	32	74.4	6070	13	ADR89183	AdR89183 Domestic
C 517	32	74.4	2016	13	ADQ89901	Adg89901 Antagonis	C 590	32	74.4	6104	4	AA546296	Aa546296 Tumour su
C 518	32	74.4	2092	13	ADV41009	Adv41009 Rat cardi	C 591	32	74.4	6104	6	ABL32297	AbL32297 Human imm
C 519	32	74.4	2138	4	ABL24720	AbL24720 Drosophill	C 592	32	74.4	6104	10	ADB54104	AdB54104 Pretreate
C 520	32	74.4	2198	10	ADC08371	AdC08371 Rice DNA	C 593	32	74.4	6104	10	ADB54232	AdB54232 Pretreate
C 521	32	74.4	2213	5	AA568904	Aa568904 DNA encod	C 594	32	74.4	6104	13	ADS89260	AdS89260 Oligonuci
C 522	32	74.4	2213	6	AA148638	Aa148638 Human ins	C 595	32	74.4	6104	13	ADS89534	AdS89534 Oligonuci
C 523	32	74.4	2236	10	ADB63772	AdB63772 Human cdN	C 596	32	74.4	6165	6	ABN97371	AbN97371 Gene #386
C 524	32	74.4	2284	10	ADG87254	AdG87254 Mouse 133	C 597	32	74.4	6165	13	ADR52960	AdR52960 Drug ther
C 525	32	74.4	2355	10	ADF76798	AdF76798 Novel hum	C 598	32	74.4	6165	14	ABE35215	AbE35215 Human Gef
C 526	32	74.4	2361	5	ABV22379	Abv22379 Human pro	C 599	32	74.4	6181	4	ABL15836	AbL15836 Drosophill
C 527	32	74.4	2361	5	ABV29264	Abv29264 Human pro	C 600	32	74.4	6245	12	ADQ97329	AdQ97329 Mouse can
C 528	32	74.4	2361	5	ABV23476	Abv23476 Human pro	C 601	32	74.4	6728	3	AAZ35248	AaZ35248 Human BMP
C 529	32	74.4	2361	5	ABV22243	Abv22243 Human pro	C 602	32	74.4	6765	10	ADE55527	AdE55527 Rat gene
C 530	32	74.4	2361	5	ABV22359	Abv22359 Human pro	C 603	32	74.4	6765	10	ADE55531	AdE55531 Rat gene
C 531	32	74.4	2361	5	ABV29334	Abv29334 Human pro	C 604	32	74.4	6947	8	ACF34533	AcF34533 Gene enco
C 532	32	74.4	2361	5	ABV20133	Abv20133 Human pro	C 605	32	74.4	6948	5	AAH81806	AaH81806 Human dif
C 533	32	74.4	2361	5	ABV22307	Abv22307 Human pro	C 606	32	74.4	6948	14	ADX07000	AdX07000 Cyclin-de
C 534	32	74.4	2361	5	ABV22255	Abv22255 Human pro	C 607	32	74.4	6996	4	AAK53025	AaK53025 Human pol
C 535	32	74.4	2361	5	ABV22295	Abv22295 Human pro	C 608	32	74.4	7183	13	ADR84440	AdR84440 Aspergill
C 536	32	74.4	2361	5	ABV22328	Abv22328 Human pro	C 609	32	74.4	7354	4	AAK52041	AaK52041 Human pol
C 537	32	74.4	2361	5	ABV23406	Abv23406 Human pro	C 610	32	74.4	7482	10	ADC27703	AdC27703 Human col
C 538	32	74.4	2361	5	ABV28128	Abv28128 Human pro	C 611	32	74.4	7985	14	ADV97786	AdV97786 cDNA sequ
C 539	32	74.4	2361	5	ABV25966	Abv25966 Human pro	C 612	32	74.4	8021	6	ABN84542	AbN84542 Human col
C 540	32	74.4	2361	5	ABV22310	Abv22310 Human pro	C 613	32	74.4	8365	2	AAZ20056	AaZ20056 Plasmodiu
C 541	32	74.4	2361	5	ABV28199	Abv28199 Human pro	C 614	32	74.4	9190	6	AAZ139555	AaZ139555 Genomic D
C 542	32	74.4	2406	5	AAF57403	AaF57403 Human p11	C 615	32	74.4	9358	4	ABL29714	AbL29714 Drosophill
C 543	32	74.4	2475	6	AB876564	Ab876564 Human bon	C 616	32	74.4	10399	2	AAV52293	AaV52293 Streptoco
C 544	32	74.4	2534	4	ABL26322	AbL26322 Drosophill	C 617	32	74.4	10456	13	ADR84320	AdR84320 Aspergill
C 545	32	74.4	2549	3	AA9494729	Aa9494729 Soybean a	C 618	32	74.4	10647	6	AA561396	Aa561396 Human gen
C 546	32	74.4	2590	8	ABT333968	AbT333968 Human pig	C 619	32	74.4	11162	4	AAK81259	AaK81259 Human imm
C 547	32	74.4	2590	10	ADC28597	AdC28597 Human CYP	C 620	32	74.4	11162	4	AAK81258	AaK81258 Human imm
C 548	32	74.4	2652	8	ACA36947	AcA36947 Prokaryot	C 621	32	74.4	13326	6	ABL33858	AbL33858 Human imm
C 549	32	74.4	2726	9	AAL60885	AaL60885 Human CYP	C 622	32	74.4	15050	6	ABL29772	AbL29772 Drosophill
C 550	32	74.4	2730	8	ABT33967	AbT33967 Human pig	C 623	32	74.4	17789	4	ABL07956	AbL07956 Drosophill
C 551	32	74.4	2799	2	AAV71085	AaV71085 PKG-green	C 624	32	74.4	26502	12	ADL18577	AdL18577 Human wil
C 552	32	74.4	2802	2	AAV71084	AaV71084 Green flu	C 625	32	74.4	27020	13	ADT77151	AdT77151 Type II d
C 553	32	74.4	2828	6	ABZ11682	AbZ11682 Human pol	C 626	32	74.4	29072	13	ADV87723	AdV87723 Streptoco
C 554	32	74.4	2935	8	ABT42910	AbT42910 Human neu	C 627	32	74.4	29072	13	ADV78976	AdV78976 Streptoco
C 555	32	74.4	2949	12	ADK41161	AdK41161 Human tum	C 628	32	74.4	31116	11	ACN44954	AcN44954 Human gen
C 556	32	74.4	2950	4	ABL25212	AbL25212 Drosophill	C 629	32	74.4	31129	6	AAZ36229	AaZ36229 Human tra
C 557	32	74.4	2957	10	ADB62337	AdB62337 Rat gene	C 630	32	74.4	31279	14	ADZ13255	AdZ13255 Human can
C 558	32	74.4	3054	6	ABN89485	AbN89485 Human hep	C 631	32	74.4	31810	11	ACN43888	AcN43888 Mouse gen
C 559	32	74.4	3085	5	ABA17022	AbA17022 Human ner	C 632	32	74.4	33403	11	ACN44824	AcN44824 Mouse gen
C 560	32	74.4	3103	12	ADQ09942	AdQ09942 Mouse NK-	C 633	32	74.4	35169	11	ACN43908	AcN43908 Mouse gen
C 561	32	74.4	3103	12	ADQ09868	AdQ09868 Mouse NK-	C 634	32	74.4	36047	10	ADB95869	AdB95869 Mouse Nfk
C 562	32	74.4	3257	4	ABL16934	AbL16934 Drosophill	C 635	32	74.4	36048	9	ADA02621	AdA02621 Mouse Nfk
C 563	32	74.4	3273	4	ABL16936	AbL16936 Drosophill	C 636	32	74.4	36048	10	ADB72359	AdB72359 Mouse Nfk
C 564	32	74.4	3495	6	ABN70606	AbN70606 Streptoco	C 637	32	74.4	39071	12	ADM97420	AdM97420 Prostata
C 565	32	74.4	3498	13	ADR83810	AdR83810 S. pyogen	C 638	32	74.4	39121	11	ACN44168	AcN44168 Mouse gen
C 566	32	74.4	3522	6	ABN67343	AbN67343 Streptoco	C 639	32	74.4	39401	14	ADZ13218	AdZ13218 Murine ca
C 567	32	74.4	3537	12	ADM44200	AdM44200 Novel hum	C 640	32	74.4	39982	8	AAD48290	AaD48290 Human enz
C 568	32	74.4	3700	13	ADS60494	AdS60494 Bacterial	C 641	32	74.4	43599	6	ABK84242	AbK84242 Human cdN
C 569	32	74.4	3740	5	AA568905	Aa568905 DNA encod	C 642	32	74.4	50000	10	ADC87689	AdC87689 Human mam
C 570	32	74.4	3740	6	AA148639	Aa148639 Human ins	C 643	32	74.4	52679	11	ACN44216	AcN44216 Mouse gen
C 571	32	74.4	3740	13	ADR26070	AdR26070 Breast ca	C 644	32	74.4	54200	11	ACN43974	AcN43974 Human gen
C 572	32	74.4	3740	13	ADZ49093	AdZ49093 Insulin s	C 645	32	74.4	54355	9	ADA02753	AdA02753 Mouse Mor
C 573	32	74.4	3902	5	ABV30226	Abv30226 Human pro	C 646	32	74.4	54355	10	ADB72491	AdB72491 Mouse Mor
C 574	32	74.4	3940	10	ADE53669	AdE53669 Human pro	C 647	32	74.4	54355	10	ADC85233	AdC85233 Mouse Mor
C 575	32	74.4	4263	13	ADR85494	AdR85494 Aspergill	C 648	32	74.4	54355	10	ADM74348	AdM74348 Murine ca
C 576	32	74.4	4313	13	ADX61752	AdX61752 Plant ful	C 649	32	74.4	58822	9	ADA02540	AdA02540 Human TCO
C 577	32	74.4	4347	12	ADK41159	AdK41159 Human tum	C 650	32	74.4	58822	10	ADB72278	AdB72278 Human TCO
C 578	32	74.4	4416	12	ADO35387	AdO35387 Novel mou	C 651	32	74.4	58822	10	ADB95788	AdB95788 Human TCO
C 579	32	74.4	4456	13	ADR84907	AdR84907 Aspergill	C 652	32	74.4	58847	12	ADQ97328	AdQ97328 Continuation (4 of
C 580	32	74.4	4607	12	ADK41158	AdK41158 Human tum	C 653	32	74.4	59838	11	ACN44982	AcN44982 Human gen
C 581	32	74.4	4822	12	ADK41160	AdK41160 Human tum	C 654	32	74.4	73634	9	ADA02615	AdA02615 Mouse Fyn
C 582	32	74.4	4872	10	ADD45160	AdD45160 Human gen	C 655	32	74.4	73634	10	ADB72353	AdB72353 Mouse Fyn
C 583	32	74.4	5000	4	AAF62416	AaF62416 A thalian	C 656	32	74.4	73634	10	ADB95863	AdB95863 Mouse Fyn
C 584	32	74.4	5219	6	ABL32768	AbL32768 Human imm	C 657	32	74.4	73702	13	ADZ13582	AdZ13582 Human can
C 585	32	74.4	5565	4	ABL29715	AbL29715 Drosophill	C 658	32	74.4	73723	14	ABD33145	AbD33145 Human can
C 586	32	74.4	5574	6	AA563339	Aa563339 Chemical	C 659	32	74.4	78539	8	ACA64942	AcA64942 Human FRA
C 587	32	74.4	6037	14	ADV14507	AdV14507 Silkworm	C 660	32	74.4	84248	3	ABQ99651	AbQ99651 Human MS4
C 588	32	74.4	6070	10	ADE15949	AdE15949 Silkworm	C 661	32	74.4	94618	3	AAF22285	AaF22285 BAC conta

662	32	74.4	96960	8	ACF62734	AcF62734 Cancer ba	735	32	74.4	319608	5	AA509301	AA509301 Human sch
663	32	74.4	96960	8	ADB20849	ADB20849 MRP1 base	c 736	32	74.4	342748	14	ADZ13793	ADZ13793 Human can
664	32	74.4	96960	10	ADB87938	ADB87938 Human UGT	c 737	31	72.1	90	6	ABS23915	ABS23915 Human gen
665	32	74.4	96960	10	ADB96921	ADB96921 Human MDR	738	31	72.1	110	14	ADX03058	ADX03058 Human pri
666	32	74.4	96960	10	ADB992112	ADB992112 Human can	739	31	72.1	110	14	ADX03057	ADX03057 Human pri
667	32	74.4	101954	13	ABD33574	ABD33574 Human can	c 740	31	72.1	121	11	ADZ42871	ADZ42871 Human gen
668	32	74.4	106416	4	ABL18718	ABL18718 Moxephil	c 741	31	72.1	182	5	ABV32783	ABV32783 Human pro
669	32	74.4	109565	13	ABD33086	ABD33086 Moxephil	c 742	31	72.1	211	4	AAJ27958	AAJ27958 Probe #17
670	32	74.4	110000	2	AAV21209_04	Continuation (5 of	c 743	31	72.1	211	4	AAJ56941	AAJ56941 Probe #25
671	32	74.4	110000	2	AAK91990_04	Continuation (5 of	c 744	31	72.1	211	4	ABA40813	ABA40813 Probe #19
672	32	74.4	110000	6	ABN71527_07	Continuation (8 of	c 745	31	72.1	211	4	AAK50918	AAK50918 Human bon
673	32	74.4	110000	6	ABQ67196_0	ABQ67196 Listeria	c 746	31	72.1	211	4	AAK24924	AAK24924 Human bra
674	32	74.4	110000	6	ABQ69245_09	Continuation (10 of	c 747	31	72.1	211	6	ABZ24429	ABZ24429 Human gen
675	32	74.4	110000	6	ABQ69245_15	Continuation (16 of	c 748	31	72.1	222	6	ABN24596	ABN24596 Human ORF
676	32	74.4	110000	6	ABQ69245_28	Continuation (25 of	c 749	31	72.1	225	9	ADZ30252	ADZ30252 DNA encod
677	32	74.4	110000	6	ABQ67197_04	Continuation (9 of	750	31	72.1	259	5	ADI68682	ADI68682 Human ova
678	32	74.4	110000	6	ABQ67195_1	Continuation (2 of	751	31	72.1	259	5	ADI75044	ADI75044 Human ova
679	32	74.4	110000	6	ABA03041_00	ABA03041 Listeria	c 752	31	72.1	280	5	ABA11750	ABA11750 Human ner
680	32	74.4	110000	6	ABA03041_15	Continuation (16 of	753	31	72.1	280	13	ACF81341	ACF81341 Human SIR
681	32	74.4	110000	6	ABA03041_18	Continuation (19 of	754	31	72.1	294	6	ABL71470	ABL71470 Corn tass
682	32	74.4	110000	8	ABX16390_4	Continuation (5 of	c 755	31	72.1	321	4	AAK60524	AAK60524 Human imm
683	32	74.4	110000	10	ADH10017_2	Continuation (3 of	c 756	31	72.1	324	3	AAK74950	AAK74950 Human ORF
684	32	74.4	110000	10	ABS56454_13	Continuation (14 of	c 757	31	72.1	344	6	ABN94653	ABN94653 Gene #115
685	32	74.4	110000	11	ACN44150_1	Continuation (2 of	c 758	31	72.1	345	14	ABE54717	ABE54717 DNA encod
686	32	74.4	110000	11	ADQ97266_2	Continuation (3 of	759	31	72.1	357	3	AAZ27586	AAZ27586 Human sec
687	32	74.4	110000	13	ABD32923_4	Continuation (5 of	c 760	31	72.1	369	4	AAH73056	AAH73056 Human cer
688	32	74.4	110000	13	ABD32911_1	Continuation (2 of	c 761	31	72.1	379	8	ABX47144	ABX47144 Bovine cer
689	32	74.4	110000	13	ADV81204_08	Continuation (9 of	c 762	31	72.1	386	4	AAH69607	AAH69607 Human cer
690	32	74.4	110000	14	ADZ12821_1	Continuation (2 of	c 763	31	72.1	391	8	ABX48559	ABX48559 Bovine ES
691	32	74.4	110000	14	ADZ45062_02	Continuation (3 of	764	31	72.1	392	5	ABV56708	ABV56708 Human pro
692	32	74.4	110000	14	ADZ45062_04	Continuation (5 of	c 765	31	72.1	406	9	ACH45811	ACH45811 Human foe
693	32	74.4	110000	14	ADZ42285_2	Continuation (3 of	c 766	31	72.1	419	9	ACH22289	ACH22289 Human adu
694	32	74.4	110000	14	AEA61120_0	AEA61120 Human Iga	c 767	31	72.1	424	12	ADP71987	ADP71987 Renal tox
695	32	74.4	110000	14	AEA61124_1	Continuation (2 of	c 768	31	72.1	432	5	ABV41708	ABV41708 Human pro
696	32	74.4	110000	14	AEA61124_2	Continuation (3 of	c 769	31	72.1	438	10	ACF70489	ACF70489 Photorhab
697	32	74.4	110000	14	ABE33175_04	Continuation (5 of	c 770	31	72.1	442	3	AAA43640	AAA43640 Human sec
698	32	74.4	116858	11	ACN44212	Continuation (5 of	771	31	72.1	443	5	ADL40281	ADL40281 Human ova
699	32	74.4	120447	12	ADQ97391	ADQ97391 Human gen	c 772	31	72.1	446	10	ADP80573	ADP80573 Leukaemia
700	32	74.4	121167	12	ADQ97733	ADQ97733 Human can	c 773	31	72.1	460	4	AAJ59486	AAJ59486 Human pol
701	32	74.4	121378	10	ABX77171	ABX77171 DNA sequ	c 774	31	72.1	467	6	ABL84515	ABL84515 Human ova
702	32	74.4	125515	10	ADL13941	ADL13941 Osteoarth	c 775	31	72.1	468	6	ABN20296	ABN20296 Human ORF
703	32	74.4	125512	6	ABN83429	ABN83429 Human tra	c 776	31	72.1	483	10	ACF72137	ACF72137 Photorhab
704	32	74.4	127098	10	ADL13649	ADL13649 Osteoarth	c 777	31	72.1	483	13	ACN60943	ACN60943 Cotton gy
705	32	74.4	127917	13	ADR52731	ADR52731 Drug ther	c 778	31	72.1	484	4	AAI61272	AAI61272 Human pol
706	32	74.4	130505	13	ABD32544	ABD32544 Human can	c 779	31	72.1	487	13	ADW11540	ADW11540 Bacillus
707	32	74.4	130877	13	ABD33104	ABD33104 Human can	c 780	31	72.1	488	5	ADL40210	ADL40210 Human ova
708	32	74.4	144460	3	AAZ93815	AAZ93815 Olfactory	c 781	31	72.1	492	13	ADQ48943	ADQ48943 Novel can
709	32	74.4	145985	12	ADQ97164	ADQ97164 Human can	c 782	31	72.1	500	6	ABZ74528	ABZ74528 Human PAP
710	32	74.4	150830	12	ADQ97260	ADQ97260 Mouse can	c 783	31	72.1	501	14	ADM82867	ADM82867 MAP3K9 ma
711	32	74.4	154799	13	ADZ36467	ADZ36467 Mouse aut	784	31	72.1	506	12	ACH78300	ACH78300 Human gen
712	32	74.4	163319	3	APZ22306	APZ22306 Arabidops	c 785	31	72.1	516	10	ABE07402	ABE07402 Novel cod
713	32	74.4	163382	13	ABD32659	ABD32659 Human can	c 786	31	72.1	518	5	ABV11638	ABV11638 Human pro
714	32	74.4	163382	13	ABD32659	ABD32659 Human can	c 787	31	72.1	519	13	ACN49125	ACN49125 Cotton pr
715	32	74.4	165097	14	ABE61167	ABE61167 Human FLJ	c 788	31	72.1	519	14	ABE54719	ABE54719 DNA encod
716	32	74.4	165097	14	ABE61167	ABE61167 Human can	c 789	31	72.1	521	6	ABSI11558	ABSI11558 Human gen
717	32	74.4	168198	12	ADQ59452	ADQ59452 Human can	c 790	31	72.1	521	10	ABE81236	ABE81236 Arabidops
718	32	74.4	169398	14	ADZ13775	ADZ13775 Human can	791	31	72.1	523	4	AAK92931	AAK92931 Human CDN
719	32	74.4	170170	10	ADL13643	ADL13643 Osteoarth	792	31	72.1	523	12	ACN29358	ACN29358 3' end of
720	32	74.4	175590	14	ADV77908	ADV77908 Human BAC	793	31	72.1	526	13	ACN50398	ACN50398 Cotton ma
721	32	74.4	175590	14	ADL13512	ADL13512 Osteoarth	c 794	31	72.1	528	4	AAI62505	AAI62505 Human bre
722	32	74.4	184951	14	ABE42739	ABE42739 L. pneumo	c 795	31	72.1	528	4	AAI02207	AAI02207 Human rep
723	32	74.4	185021	14	ABE39169	ABE39169 Human gen	c 796	31	72.1	539	5	ABV02469	ABV02469 Human pro
724	32	74.4	186391	11	ACN43938	ACN43938 Human can	c 797	31	72.1	540	10	ABX07002	ABX07002 S. pneumo
725	32	74.4	186854	14	ADZ38909	ADZ38909 Vaccinia	c 798	31	72.1	543	3	AAA05545	AAA05545 Streptoco
726	32	74.4	194588	14	ABE35717	ABE35717 L. pneumo	c 799	31	72.1	543	13	ADK45340	ADK45340 Streptoco
727	32	74.4	200000	12	ADQ47192	ADQ47192 DNA sequ	c 800	31	72.1	544	13	ACN57075	ACN57075 Cotton gy
728	32	74.4	236303	4	AA511614	AA511614 Human gen	801	31	72.1	549	6	ABN62345	ABN62345 Human can
729	32	74.4	240000	8	ACD13446	ACD13446 Human DNA	802	31	72.1	550	4	AAK92517	AAK92517 Human CDN
730	32	74.4	252907	13	ABD32694	ABD32694 Human can	803	31	72.1	550	12	ADL28944	ADL28944 3' end of
731	32	74.4	271990	10	ADD25213	ADD25213 Fertility	c 804	31	72.1	554	4	AAI19129	AAI19129 Probe #90
732	32	74.4	271990	12	ADN61228	ADN61228 Radish nu	c 805	31	72.1	554	4	AAI44278	AAI44278 Probe #12
733	32	74.4	273254	3	AAK81914	AAK81914 Chlamydia	c 806	31	72.1	554	4	ABA31291	ABA31291 Probe #97
734	32	74.4	319608	3	AAH51601	AAH51601 Human chr	c 807	31	72.1	554	4	AAK38328	AAK38328 Human bon

C 808	31	72.1	554	4	AAX12615	Aak12615 Human bra	881	31	72.1	860	11	ACL32803	ACL32803 Rice abio
C 809	31	72.1	554	6	ABS12382	Abel12382 Human gen	C 882	31	72.1	863	11	ACN83701	ACN83701 Breast ca
C 810	31	72.1	558	8	ACF74327	Acfr74327 Staphyloc	883	31	72.1	870	8	ABZ42344	ABZ42344 Streptoco
C 811	31	72.1	558	9	ACH27156	Ach27156 Human adu	884	31	72.1	870	13	ADR92236	ADR92236 Novel S.
C 812	31	72.1	566	6	ABN62672	Abn62672 Human can	885	31	72.1	870	14	AEA56106	AEA56106 Streptoco
C 813	31	72.1	568	4	AAI84599	Aai84599 Human pol	C 886	31	72.1	903	4	AAG63550	Aag63550 Human imm
C 814	31	72.1	573	10	ABZ84713	Abz84713 Toxicolog	C 887	31	72.1	911	13	ADX48924	Adx48924 Plant ful
C 815	31	72.1	576	5	ABV54882	Abv54882 Human pro	888	31	72.1	934	8	ACC78301	Acc78301 DNA encod
C 816	31	72.1	578	5	ABF12092	Abf12092 Aspergill	889	31	72.1	960	14	AEA26768	AEA26768 Stress to
C 817	31	72.1	578	5	ABV11039	Abv11039 Human pro	C 890	31	72.1	966	3	AAC46230	Aac46230 Arabidops
C 818	31	72.1	578	13	ADU56133	Adu56133 Aspergill	891	31	72.1	966	6	ABN69282	Abn69282 Streptoco
C 819	31	72.1	578	14	ADZ94136	Adz94136 Aspergill	C 892	31	72.1	969	3	AAC35676	Aac35676 Arabidops
C 820	31	72.1	600	6	ABN92309	Abn92309 Staphyloc	893	31	72.1	969	3	AAC48346	Aac48346 Arabidops
C 821	31	72.1	600	13	ADS03918	Ads03918 Staphyloc	894	31	72.1	969	13	ADU69316	Adu69316 S agalact
C 822	31	72.1	601	6	ABQ58768	Abq58768 Human col	895	31	72.1	969	13	ADV83455	Adv83455 Streptoco
C 823	31	72.1	601	13	ADU47779	Adu47779 Human tra	C 896	31	72.1	971	14	ADZ13154	Adz13154 Human can
C 824	31	72.1	601	14	ACL59155	ACL59155 Human col	897	31	72.1	972	8	ACL22008	ACL22008 Prokaryot
C 825	31	72.1	610	4	AAS61041	Aas61041 Human can	C 898	31	72.1	980	4	AAD06338	Aad06338 Pentactlet
C 826	31	72.1	617	5	ABV54245	Abv54245 Human pro	899	31	72.1	984	3	AAC39467	Aac39467 Arabidops
C 827	31	72.1	618	5	ABV55693	Abv55693 Human pro	900	31	72.1	984	3	AAC48347	Aac48347 Arabidops
C 828	31	72.1	624	8	ACA18760	Acad18760 Prokaryot	901	31	72.1	988	6	ABN74648	Abn74648 Bovine em
C 829	31	72.1	634	12	ADN11398	Adn11398 Human pro	902	31	72.1	992	3	AAC33056	Aac33056 Arabidops
C 830	31	72.1	635	5	ABV01870	Abv01870 Human pro	903	31	72.1	1000	9	ADB23202	Adb23202 Envinrome
C 831	31	72.1	642	9	ADA88788	Ada88788 Streptoco	904	31	72.1	1001	4	AAC88189	Aac88189 Optimum p
C 832	31	72.1	650	14	AEA20755	Aea20755 Novel hum	905	31	72.1	1024	3	AAA09239	Aaa09239 Mastrevir
C 833	31	72.1	667	4	AAL22500	Aal22500 Human bre	C 906	31	72.1	1045	12	ADP22487	Adp22487 Sea-equir
C 834	31	72.1	667	6	ADG79323	Adg79323 Human sec	907	31	72.1	1052	3	AAA73225	Aaa73225 Human V-t
C 835	31	72.1	674	4	AAX80237	Aax80237 Human imm	908	31	72.1	1054	4	ABL111273	Ab111273 Drosophil
C 836	31	72.1	674	4	AAX80236	Aax80236 Human imm	909	31	72.1	1055	11	ACL36658	ACL36658 Rice stre
C 837	31	72.1	684	5	ABV41118	Abv41118 Human pro	C 910	31	72.1	1080	9	ADA88796	Ada88796 Streptoco
C 838	31	72.1	684	5	ABV32187	Abv32187 Human pro	911	31	72.1	1081	8	ACA57510	Aca57510 Human adi
C 839	31	72.1	686	5	ABV22927	Abv22927 Human pro	912	31	72.1	1090	3	AAC40430	Aac40430 Arabidops
C 840	31	72.1	686	5	ABV28758	Abv28758 Human pro	913	31	72.1	1110	3	AAA09238	Aaa09238 Geminivir
C 841	31	72.1	686	13	ADQ48942	Adq48942 Novel can	914	31	72.1	1112	10	ADC20846	Adc20846 Human sec
C 842	31	72.1	696	6	ABT09200	Abt09200 Phase-1 R	915	31	72.1	1112	10	ABT16983	Abt16983 Human sec
C 843	31	72.1	696	10	ADG30974	Adg30974 Liver tox	916	31	72.1	1112	10	ABZ67931	Abz67931 Human sec
C 844	31	72.1	696	12	ADG45593	Adg45593 Liver inf	917	31	72.1	1119	6	ABK76999	Abk76999 Bacillus
C 845	31	72.1	696	13	ADR91217	Adr91217 Spleen ne	C 918	31	72.1	1147	4	AAF54994	Aaf54994 Nucleotid
C 846	31	72.1	697	4	AAI96695	Aai96695 Human neu	919	31	72.1	1149	12	ADN05709	Adn05709 Antipepti
C 847	31	72.1	700	2	ADR01610	Adr01610 A. gossyp	920	31	72.1	1160	3	AAF22348	Aaf22348 Human sec
C 848	31	72.1	705	6	ABG64986	Abg64986 Invertebr	921	31	72.1	1160	10	ADC20223	Adc20223 Human sec
C 849	31	72.1	708	6	ABK73211	Abk73211 Bacillus	922	31	72.1	1160	10	ABT16829	Abt16829 Human sec
C 850	31	72.1	710	4	AAH70347	Aah70347 Human cer	923	31	72.1	1160	10	ABZ67143	Abz67143 Human sec
C 851	31	72.1	714	2	ADR02171	Adr02171 A. gossyp	C 924	31	72.1	1160	14	ADY66015	Ady66015 S. mansoni
C 852	31	72.1	714	2	ADR02447	Adr02447 A. gossyp	C 925	31	72.1	1164	5	AAD08105	Aad08105 Novel hum
C 853	31	72.1	714	4	AAL13631	Aal13631 Human bre	C 926	31	72.1	1173	10	ADB87361	Adb87361 Transgene
C 854	31	72.1	715	4	ABL13821	Ab13821 Drosophil	C 927	31	72.1	1173	12	ADJ35170	Adj35170 Human sta
C 855	31	72.1	716	12	ADJ43437	Adj43437 Plant cDN	C 928	31	72.1	1173	12	ADL96788	Adl96788 Stablizin
C 856	31	72.1	722	2	ADR01988	Adr01988 A. gossyp	C 929	31	72.1	1176	12	ADQ25939	Adq25939 Protein p
C 857	31	72.1	723	4	AAH08374	Aah08374 Human cDN	C 930	31	72.1	1176	12	ADQ28708	Adq28708 STAR elem
C 858	31	72.1	725	2	ADR02359	Adr02359 A. gossyp	C 931	31	72.1	1176	14	AEA51516	Aea51516 Stablizi
C 859	31	72.1	727	4	AAH05135	Aah05135 Human cDN	C 932	31	72.1	1181	5	AAD08103	Aad08103 Novel hum
C 860	31	72.1	728	3	AAH01788	Aah01788 Human col	933	31	72.1	1199	3	AAC46231	Aac46231 Arabidops
C 861	31	72.1	729	6	ABK73086	Abk73086 Bacillus	934	31	72.1	1201	3	AAC35313	Aac35313 Arabidops
C 862	31	72.1	732	8	ACC78302	Acc78302 DNA encod	C 935	31	72.1	1225	10	ADC87608	Adc87608 Human GPC
C 863	31	72.1	734	5	AAS92615	Aas92615 DNA encod	C 936	31	72.1	1230	2	AAX91111	Aax91111 Group B S
C 864	31	72.1	747	4	AAI96802	Aai96802 Human neu	937	31	72.1	1238	10	ADC20845	Adc20845 Human sec
C 865	31	72.1	754	2	AAZ34492	Aaz34492 Human LYS	938	31	72.1	1238	10	ABT16982	Abt16982 Human sec
C 866	31	72.1	754	6	ABG76460	Abg76460 cDNA enco	939	31	72.1	1238	10	ABZ67930	Abz67930 Human sec
C 867	31	72.1	754	6	ABX04184	Abx04184 Human mRN	940	31	72.1	1239	6	ABA96107	Ab96107 Schizosac
C 868	31	72.1	761	2	ADR01690	Adr01690 A. gossyp	941	31	72.1	1278	4	ABL24433	Ab124433 Drosophil
C 869	31	72.1	773	2	AAH40998	Aah40998 DNA encod	942	31	72.1	1278	5	AAS21831	Aas21831 Human col
C 870	31	72.1	773	2	AAV53463	Aav53463 DNA encod	943	31	72.1	1282	2	AAX12967	Aax12967 Enterococ
C 871	31	72.1	789	8	ACA39311	Ac39311 Prokaryot	944	31	72.1	1282	6	ABS98762	Abs98762 Enterococ
C 872	31	72.1	795	14	AEC02063	Aec02063 Nucleotid	C 945	31	72.1	1301	12	ADD35593	Add35593 S. agalac
C 873	31	72.1	804	10	ABX07989	Abx07989 S. pneumo	C 946	31	72.1	1302	6	ABN66519	Abn66519 Streptoco
C 874	31	72.1	804	12	ADM92030	Adm92030 S. pneumon	947	31	72.1	1305	2	AAX37246	Aax37246 Human 3-O
C 875	31	72.1	807	13	ADK45128	Adk45128 Streptoco	948	31	72.1	1305	2	AAX91110	Aax91110 Group B S
C 876	31	72.1	826	3	AAC59685	Aac59685 Human sec	949	31	72.1	1305	3	AZ36687	Aaz36687 Nucleotid
C 877	31	72.1	840	6	ABQ68553	Abq68553 Listeria	950	31	72.1	1305	8	AAL54122	Aal54122 Human 3-O
C 878	31	72.1	840	10	ADP02698	Adp02698 Bacterial	951	31	72.1	1305	10	ADD18953	Add18953 Human dis
C 879	31	72.1	851	5	ADI74973	Adi74973 Human ova	952	31	72.1	1305	12	ADJ74806	Adj74806 Marker ge
C 880	31	72.1	851	5	ADI68609	Adi68609 Human ova	953	31	72.1	1305	12	ADL82932	Adl82932 Human PRO

c 954 31 72.1 1305 12 ADO39561 Streptoco
 955 31 72.1 1305 13 ADP23866 PRO poly
 956 31 72.1 1305 13 ADP23866 PRO poly
 c 957 31 72.1 1305 13 ADU69309 S agalact
 c 958 31 72.1 1305 13 ADU69309 S agalact
 959 31 72.1 1309 4 AHI13964 Human cDN
 c 960 31 72.1 1311 10 ACG61770 Gene sequ
 c 961 31 72.1 1311 10 ACG61770 Gene sequ
 c 962 31 72.1 1323 6 ABZ32067 Candida a
 c 963 31 72.1 1344 5 AD08104 Novel hum
 964 31 72.1 1347 4 AHI53173 S. epider
 c 965 31 72.1 1348 3 AAZ65288 Human sec
 c 966 31 72.1 1348 10 ADE11677 Human sec
 c 967 31 72.1 1361 5 AD08102 Novel hum
 c 968 31 72.1 1376 6 ABN70313 Streptoco
 c 969 31 72.1 1376 12 ADK99662 Streptoco
 c 970 31 72.1 1376 12 AD010478 DNA encod
 c 971 31 72.1 1376 14 ADZ02916 S. agalac
 c 972 31 72.1 1386 3 AA61267 Human sec
 c 973 31 72.1 1389 10 ACC61872 Gene sequ
 c 974 31 72.1 1389 10 ADK64985 Disease t
 c 975 31 72.1 1390 11 AD892491 B. lichen
 c 976 31 72.1 1395 13 AD858270 Bacterial
 c 977 31 72.1 1403 3 AA73326 Human V-t
 978 31 72.1 1404 2 AAX90185 Mouse pre
 979 31 72.1 1407 10 ADB58337 Toxicity-
 980 31 72.1 1407 10 ADB58337 Primary r
 c 981 31 72.1 1416 2 AAQ96290 Partial s
 c 982 31 72.1 1416 2 AAQ96290 Partial s
 c 983 31 72.1 1416 3 AAZ46728 Bovine al
 c 984 31 72.1 1416 5 AAD16399 Bovine al
 985 31 72.1 1428 3 AAC50619 Arabidops
 c 986 31 72.1 1491 4 ABL05983 Drosophil
 c 987 31 72.1 1521 5 AAD08095 Novel hum
 c 988 31 72.1 1521 5 AAD08095 Novel hum
 c 989 31 72.1 1538 14 ADZ13150 Human can
 c 990 31 72.1 1539 5 AD08111 Novel hum
 c 991 31 72.1 1548 10 ACF68040 Photorhab
 c 992 31 72.1 1566 11 ACN44779 Human mRN
 c 993 31 72.1 1566 14 ADZ13156 Human can
 c 994 31 72.1 1572 6 ABZ17337 Arabidops
 c 995 31 72.1 1573 3 AAC67654 Human sec
 c 996 31 72.1 1578 3 AA08685 PMWAV-1 R
 c 997 31 72.1 1589 10 ADD44828 Human gen
 998 31 72.1 1592 6 ABQ54504 Human ova
 999 31 72.1 1614 8 ACC78305 DNA encod
 1000 31 72.1 1614 8 ACF12946 Human cer

ALIGNMENTS

RESULT 1
 ABN56274
 ID ABN56274 standard; DNA; 65 BP.
 XX
 AC ABN56274;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:29022.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Mus musculus.
 XX
 XX WO200210449-A2.
 XX
 XX 07-FEB-2002.
 XX
 XX 20-JUL-2001; 2001WO-IB001903.
 XX
 XX 28-JUL-2000; 2000US-0221607P.
 PR

PR 02-MAY-2001; 2001US-0287724P.
 (COMP-) COMPUGEN INC.
 PA Shoshan A, Wasserman A, Mintz E, Mintz L, Paigler S;
 PI WPI; 2002-257383/30.
 XX
 XX New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of a
 PT genome, useful for detecting tissue-, pathology-, and developmental-
 PT specific genes.
 XX
 XX Example 1; SEQ ID NO 29022; 47pp; English.
 PS
 XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
 CC)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises several
 CC oligonucleotides, each capable of hybridising selectively to a set of
 CC messenger RNAs transcribed from a given transcription unit of the genome,
 CC which encodes one or more messenger RNA splice variants. The
 CC oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterising the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a particular
 CC biological or pathological state, and so allowing the detection of tissue
 CC - and pathology-specific genes such as those genes only expressed in
 CC a specific tissue under a specific pathological condition; to detect
 CC developmental specific genes; and to detect RNA transcripts and splice
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the
 CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 65 BP; 9 A; 21 C; 12 G; 23 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.783 Length: 65
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x ABN56274 (1-65)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 29 TCCCTGCAGACTTCTTATGCTTCTTA 55

RESULT 2
 ACD97670
 ID ACD97670 standard; cDNA; 108 BP.
 XX
 AC ACD97670;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human colon cancer cell expressed cDNA #6082.
 XX

KW Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KW agriculture; food crop genome; resistance gene; retrovirus;
 KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.
 XX
 XX Homo sapiens.
 OS

XX PN US2002155438-A1.
 XX PD 24-OCT-2002.
 XX PF 27-SEP-1999; 99US-00406117.
 XX PR 20-NOV-1998; 98US-00196716.
 XX PA (SIMP/) SIMPSON A J G.
 XX PA (NETO/) NETO E D.
 XX PA (BREN/) BRENTANI R R.
 XX PI Simpson AJG, Neto ED, Brentani RR;
 XX WPI; 2003-182626/18.
 XX PT Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 XX PS Example 9; Page 866; 959pp; English.
 XX CC The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (I) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. Further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a foetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 XX SQ Sequence 108 BP; 18 A; 33 C; 22 G; 35 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.39 Length: 108
 Score: 43.00 Matches: 108
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-7 (1-9) x ACD97670 (1-108)
 QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 Db 60 TCCCTGCAACCTCTTATGCTTCTG 86

RESULT 3
 AAS87174
 ID AAS87174 standard; cDNA; 453 BP.
 XX AC AAS87174;
 XX DT 13-FEB-2002 (first entry)
 XX DE DNA encoding novel human diagnostic protein #22978.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 XX PR 23-AUG-2000; 2000US-00649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PT Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX P-PSDB; ABG22987.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 1; SEQ ID NO 22978; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 453 BP; 108 A; 111 C; 113 G; 121 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.99 Length: 453
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0
 US-10-774-176-7 (1-9) x AAS87174 (1-453)

Oy 1 SerLeuGlnThrSerTyValpHeLeu 9
Db 238 TCCCTGCAACCTCTTATGTCCTG 264

RESULT 4
ADU11677
ID ADU11677 standard; DNA; 475 BP.
XX
AC ADU11677;
XX
DT 27-JAN-2005 (first entry)
XX
DE Solid tumour prognosis gene seqid 2116.
XX
KW cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBM; prognosis; ds.
XX
OS Unidentified.
XX
PN WO2004097052-A2.
XX
XX 11-NOV-2004.
XX
XX 29-APR-2004; 2004WO-US013587.
XX
XX 29-APR-2003; 2003US-0466067P.
XX
XX 23-JAN-2004; 2004US-0538246P.
XX
XX (AMHP) WYETH.
XX
XX (STRA) STRAHS A.
XX
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
XX Immerman F, Dorner AJ;
XX
XX WPI; 2004-804779/79.
XX
XX A method, useful for prognosing and treating solid tumor, comprises
XX comparing an expression profile of a gene expressed in peripheral blood
XX mononuclear cells to a reference expression profile of a gene.
XX
XX Disclosure; Page; 111pp; English.

The invention describes a method comprising comparing an expression
profile of at least one gene in a peripheral blood sample of a patient to
at least one reference expression profile of the at least one gene, where
the patient has a solid tumour, and each of the gene is differentially
expressed in peripheral blood mononuclear cells (PBMCs) of a first class
of patients as compared to PBMCs of a second class of patients, where
both the first and second classes of patients have the solid tumour, and
each of the first and second classes is a subcluster formed by an
unsupervised clustering analysis of gene expression profiles in PBMCs of
a population of patients who have the solid tumour, and where the
majority of the first class of patients has a first clinical outcome, and
the majority of the second class of patients has a second clinical
outcome. Also described are: a system comprising (i) a memory or a
storage medium including data that represent an expression profile of at
least one gene in a peripheral blood sample of a patient who has a solid
tumour, (ii) at least another storage medium including data that
represent at least one reference expression profile of the gene, (iii) a
program capable of comparing the expression profile to the reference
expression profile, and (iv) a processor capable of executing the
program, where expression levels of the gene in peripheral blood
mononuclear cells of patients who have the solid tumour correlate with
clinical outcomes of the patients; and a nucleic acid or protein array
comprising concentrated probes for solid tumour prognosis genes, where
each of the solid tumour prognosis genes is differentially expressed in
PBMCs of a first class of patients as compared to PBMCs of a second class
of patients, where both the first and second classes of patients have a
solid tumour, and where the first class of patients has a first clinical
outcome, and the second class of patients has a second clinical outcome.
The method, system, and array are useful for prognosing and treating
solid tumours. This sequence represents a solid tumour prognosis gene of
the invention. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores: 7.38 Length: 475
Pred. No.: 43.00 Matches: 9
Score: 43.00
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-7 (1-9) x ADU11677 (1-475)

Oy 1 SerLeuGlnThrSerTyValpHeLeu 9
Db 330 TCCCTGCAACCTCTTATGTCCTG 356

RESULT 5
AAA27060
ID AAA27060 standard; DNA; 901 BP.
XX
AC AAA27060;
XX
DT 22-AUG-2000 (first entry)
XX
DE Canine 5T4 tumour-associated antigen gene.
XX
KW Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
OS Canis sp.
XX

Key Location/Qualifiers
FH 1..858
FT CDS
FT /tag= a
FT /product= "5T4 antigen"
FT 61..74
FT /tag= b
FT /note= "given in the specification but does not seem to
be part of the coding sequence and does not encode any
corresponding amino acids"
FT 135..146
FT /tag= c
FT /note= "given in the specification but does not seem to
be part of the coding sequence and does not encode any
corresponding amino acids"
FT 207..216
FT /tag= d
FT /note= "given in the specification but does not seem to
be part of the coding sequence and does not encode any
corresponding amino acids"
FT 277..290
FT /tag= e
FT /note= "given in the specification but does not seem to
be part of the coding sequence and does not encode any
corresponding amino acids"
FT 351..361
FT /tag= f
FT /note= "given in the specification but does not seem to
be part of the coding sequence and does not encode any
corresponding amino acids"
FT 422..436
FT /tag= g
FT /note= "given in the specification but does not seem to
be part of the coding sequence and does not encode any
corresponding amino acids"
FT 497..511
FT /tag= h
FT /note= "given in the specification but does not seem to
be part of the coding sequence and does not encode any

RESULT 7
 ABX76333
 ID ABX76333 standard; DNA; 927 BP.
 XX AC
 XX ABX76333;
 DT 02-APR-2003 (first entry)
 XX
 XX Lung cancer-associated polynucleotide #197.
 XX
 XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 XX Unidentified.
 XX OS
 XX
 XX W0200286443-A2.
 XX
 XX 31-OCT-2002.
 XX
 XX 18-APR-2002; 2002WO-US012476.
 XX
 XX 18-APR-2001; 2001US-0284770P.
 XX
 XX 10-MAY-2001; 2001US-0290492P.
 XX
 XX 09-NOV-2001; 2001US-0339245P.
 XX
 XX 13-NOV-2001; 2001US-0350666P.
 XX
 XX 29-NOV-2001; 2001US-0334370P.
 XX
 XX 12-APR-2002; 2002US-0372246P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PA
 XX
 XX Aziz N, Murray R;
 XX
 XX WPI; 2003-093161/08.
 XX
 XX P-PSDB; ABU56604.
 XX
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 XX for treating lung cancer, by contacting a biological sample from the
 XX patient with a polynucleotide that exhibits increased or decreased
 XX expression in lung cancer.
 XX
 XX Claim 22; Page 336; 453pp; English.
 XX PS
 XX
 XX The invention relates to a method for detecting a lung cancer-associated
 XX transcript in a cell from a patient, comprising contacting a biological
 XX sample from the patient with a polynucleotide that selectively hybridizes
 XX to a sequence that is at least 80 % identical to a gene that exhibits
 XX increased or decreased expression in lung cancer samples. Lung cancer-
 XX associated polynucleotides and polypeptides are used for identifying a
 XX compound that modulates a lung cancer-associated polypeptide, for
 XX inhibiting proliferation of a lung cancer-associated cell to treat lung
 XX cancer in a patient and for treating a mammal having lung cancer by
 XX administering a modulatory compound identified. The methods are useful
 XX for treating lung cancer, such as small cell lung cancer, non-small cell
 XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 XX emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 XX hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 XX bronchiectasis. The genes, polynucleotides and polypeptides are useful
 XX for diagnostic purposes and as targets for screening for therapeutic
 XX compounds that modulate lung cancer, such as antibodies. Sequences
 XX ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 XX invention
 XX
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 15.7 Length: 927
 XX Score: 43.00 Matches: 9
 XX Percent Similarity: 100.0% Conservative: 0
 XX Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-7 (1-9) x ABX76333 (1-927)
 QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 DB 709 TCCCTGCAACCTCTTATGTCTTCCTG 735
 RESULT 8
 ADB80503
 ID ADB80503 standard; DNA; 927 BP.
 XX AC
 XX ADB80503;
 XX
 XX 04-DEC-2003 (first entry)
 XX
 XX Ovarian cancer-associated transcript #34.
 XX
 XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 XX post-operative chemotherapy; radiation therapy; tumour prognosis;
 XX pre-cancerous lesion detection; ds; gene.
 XX
 XX Homo sapiens.
 XX OS
 XX
 XX Key Location/Qualifiers
 XX CDS 1..927
 XX /*tag= a
 XX
 XX W02002102235-A2.
 XX
 XX 27-DEC-2002.
 XX
 XX 18-JUN-2002; 2002WO-US019297.
 XX
 XX 18-JUN-2001; 2001US-0299234P.
 XX
 XX 27-AUG-2001; 2001US-0315287P.
 XX
 XX 05-SEP-2001; 2001US-0317544P.
 XX
 XX 13-NOV-2001; 2001US-0350666P.
 XX
 XX 12-APR-2002; 2002US-0372246P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PA
 XX
 XX Mack DH, Gish KC;
 XX
 XX WPI; 2003-167431/16.
 XX
 XX P-PSDB; ADB80504.
 XX
 XX Detecting an ovarian cancer-associated transcript in a cell from a
 XX patient, comprises contacting a biological sample from the patient with a
 XX polynucleotide that hybridizes to an ovarian cancer gene.
 XX
 XX Claim 10; Page 297; 332pp; English.
 XX PS
 XX The invention relates to a method of detecting an ovarian cancer-
 XX associated transcript in a cell from a patient, by contacting a
 XX biological sample from the patient with a polynucleotide that selectively
 XX hybridizes to a sequence at least 80% identical to any of one of 80
 XX nucleic acid sequences given in the specification. The method is useful
 XX in diagnosing ovarian cancer and in identifying and using agents and/or
 XX targets that inhibit ovarian cancer. The nucleic acid molecule,
 XX polypeptide and the antibody may also be used in detecting ovarian
 XX cancers, monitoring and early detection of relapse following treatment,
 XX monitoring response to therapy, selecting patients for post-operative
 XX chemotherapy or radiation therapy, in selecting mode of therapy,
 XX determining tumour prognosis, early detection of pre-cancerous lesions,
 XX and as vaccines. This sequence corresponds to one of the nucleic acids
 XX used for the detection method of the invention.
 XX
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 15.7 Length: 927
 XX Score: 43.00 Matches: 9
 XX Percent Similarity: 100.0% Conservative: 0
 XX Best Local Similarity: 100.0% Mismatches: 0

```

Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ADB80503 (1-927)
QY 1 SerLeuGlnThrSerTyrValPheLeu 9
DB 709 TCCCTGCAAACTCTTATGTCTTCCTG 735

RESULT 9
ID ADN38723
ID ADN38723 standard; cDNA; 927 BP.
XX
AC ADN38723;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX PD 22-MAY-2003.
XX
XX PF 13-NOV-2002; 2002WO-US036810.
XX
XX PR 13-NOV-2001; 2001US-0350666P.
XX
XX PR 21-NOV-2001; 2001US-0332464P.
XX
XX PR 29-NOV-2001; 2001US-0334393P.
XX
XX PR 03-DEC-2001; 2001US-0335394P.
XX
XX PR 14-DEC-2001; 2001US-034376P.
XX
XX PR 08-JAN-2002; 2002US-0347211P.
XX
XX PR 10-JAN-2002; 2002US-0347349P.
XX
XX PR 08-FEB-2002; 2002US-0355250P.
XX
XX PR 13-FEB-2002; 2002US-0356714P.
XX
XX PR 20-FEB-2002; 2002US-0359077P.
XX
XX PR 29-MAR-2002; 2002US-0368809P.
XX
XX PR 04-APR-2002; 2002US-0370110P.
XX
XX PR 12-APR-2002; 2002US-0372246P.
XX
XX PR 05-JUN-2002; 2002US-0386614P.
XX
XX PR 16-JUL-2002; 2002US-0396839P.
XX
XX PR 22-JUL-2002; 2002US-0397775P.
XX
XX PR 22-JUL-2002; 2002US-0397845P.
XX
XX PR 09-SEP-2002; 2002US-0409450P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX AFAR D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevesi PA;
XX Mack DR, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
XX P-PSDB; ADN38724.
XX
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX
XX Claim 8; SEQ ID NO 41; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX whose expression is upregulated or downregulated in specific cancers or

```

```

CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides, and
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.: 15.7 Length: 927
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-7 (1-9) x ADN38723 (1-927)
QY 1 SerLeuGlnThrSerTyrValPheLeu 9
DB 709 TCCCTGCAAACTCTTATGTCTTCCTG 735

RESULT 10
AAD56198
ID AAD56198 standard; DNA; 973 BP.
XX
XX AC AAD56198;
XX
XX DT 07-AUG-2003 (first entry)
XX
XX DE Human LRRCAPS related DNA #5.
XX
XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
XX LRRCAPS; cancer; gene therapy; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2003035831-A2.
XX
XX PD 01-MAY-2003.
XX
XX PF 21-OCT-2002; 2002WO-US033540.
XX
XX PR 22-OCT-2001; 2001US-0338733P.
XX
XX PR 15-FEB-2002; 2002US-0357600P.
XX
XX PR 01-MAR-2002; 2002US-0361196P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
XX Francis-Lang H, Friedman L;
XX WPI; 2003-421410/39.
XX
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
XX comprises contacting an assay system comprising a purified leucine rich
XX repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
XX Example 5; Page 74-75; 99pp; English.
XX
XX The invention relates to a method of identifying a candidate p53 pathway
XX modulating agent. The method involves contacting an assay system
XX comprising a purified Leucine rich repeat, capricious related (LRRCAPS)

```

CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX

SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 16.6 Length: 973
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x AAD56198 (1-973)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 724 TCCTGCAACCTCTTATGTCCTCTG 750

RESULT 11

ABV99349

ID ABV99349 standard; DNA; 1156 BP.

AC ABV99349;

DT 27-JAN-2003 (first entry)

XX Human NOV8a coding sequence.

DE Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 XX antinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
 KW notropic; immunosuppressive; osteoprotic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

XX Homo sapiens.

OS WO200272771-A2.

PN 19-SEP-2002.

PD 08-MAR-2002; 2002WO-US007288.

PF 08-MAR-2001; 2001US-0274101P.

PP 08-MAR-2001; 2001US-0274194P.

XX 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274322P.

PR 09-MAR-2001; 2001US-0274849P.

PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.

PR 13-MAR-2001; 2001US-0275579P.

PR 13-MAR-2001; 2001US-0275601P.

PR 16-MAR-2001; 2001US-0276000P.

PR 19-MAR-2001; 2001US-0276994P.

PR 20-MAR-2001; 2001US-0277239P.

PR 20-MAR-2001; 2001US-0277321P.

PR 27-MAR-2001; 2001US-0278999P.
 PR 28-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0279995P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 02-APR-2001; 2001US-0280900P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 03-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 15-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 30-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 21-NOV-2001; 2001US-0333272P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337428P.
 PR 03-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.

(CURA-) CURAGEN CORP.

Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;

Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;

Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;

Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;

Taupier RJ, Fadigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PP;

Zhong M;

WPI; 2002-732824/79.

P-PSDB; ABP70071.

New NOVX polypeptides and polynucleotides, useful for preventing, diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer, Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic disorders, and asthma.

Claim 16; Page 114-115; 619pp; English.

The present invention relates to new isolated proteins (NOVX) and their coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is any number from 1 to 48. The NOVX proteins and coding sequences are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder. The NOVX coding sequences and proteins are useful for treating, diabetes, preventing or diagnosing diseases such as metabolic disorders, cancer, cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, cardiovascular

CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substructures for use in
 CC therapeutic or diagnostic methods

XX SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 20.1 Length: 1156
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x ABV99349 (1-1156)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 DB 940 TCCTGCAACCTCTTATGCTTCTG 966

RESULT 12

ID ABK87175
 AC ABK87175 standard; cDNA; 1260 BP.

XX AC ABK87175;

XX DT 07-OCT-2002 (first entry)

XX DE cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.

XX KW Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.

XX OS Felis sp.

XX FH Key Location/Qualifiers
 XX CDS 1..1260
 XX FT /*tag= a
 XX FT /product= "5T4 protein"

XX FN WO200238612-A2.

XX PD 16-MAY-2002.

XX PP 13-NOV-2001; 2001WO-GB005004.

XX PR 13-NOV-2000; 2000WO-GB004317.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI Myers K, Drury N, Carroll M;

XX XX WPI; 2002-557449/59.

DR P-PSDB; AAU98694.

XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.

XX PS Claim 4; Page 68; 68pp; English.

XX CC The present invention relates to the isolation of canine and feline

CC oncofoetal leucine-rich glycoproteins known as 5T4, and the

CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are

CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 22.2 Length: 1260
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x ABK87175 (1-1260)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 DB 1048 TCCTGCAGACTTCTTATGCTTCTTA 1074

RESULT 13

ADB97513

ID ADB97513 standard; DNA; 1260 BP.

XX AC ADB97513;

XX DT 04-DEC-2003 (first entry)

XX DE Feline 5T4 antigen DNA.

XX KW Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX CDS 1..1260
 XX FT /*tag= a
 XX FT /product= "Feline 5T4 antigen protein"

XX PN WO2003068816-A1.

XX PD 21-AUG-2003.

XX PP 13-FEB-2003; 2003WO-GB000670.

XX PR 13-FEB-2002; 2002GB-00003419.

XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX PI Carroll M, Kingsman S, Redchenko I;

XX XX WPI; 2003-637141/60.

DR P-PSDB; ADB97520.

XX New major histocompatibility complex class I peptide epitopes from human
 PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.

XX PS Disclosure; Page 67; 73pp; English.

XX CC The invention relates to a novel Major Histocompatibility Complex (MHC)

CC class I peptide epitope of the 5T4 antigen. The invention further
 CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 CC epitope, a vector system capable of delivering the 5T4 epitope nucleic
 CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
 CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 5T4 epitope and/or its encoding nucleic acid; a method
 CC comprising detecting the presence of the 5T4 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 5T4 epitope in conjunction with an MHC class
 CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC particularly cancer. The detection method is useful for diagnosing or
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline
 CC 5T4 antigen coding DNA of the invention.

SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 22.2 Length: 1260
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ADB97513 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 1048 TCCCTGCAGACTCTTATGCTTCTTCTA 1074

RESULT 14

ADB97452

ID ADB97452 standard; DNA; 1260 BP.

XX AC ADB97452;

XX 04-DEC-2003 (first entry)

XX DNA encoding feline 5T4 protein.

XX gene, ds; feline; Major Histocompatibility Complex class II; MHC;

XX epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.

XX Unidentified.

XX Key Location/Qualifiers

XX CDS 1..1260

XX /*tag= a

XX /product= "Feline 5T4 antigen protein"

XX WO2003068815-A2.

XX 21-AUG-2003.

XX 13-FEB-2003; 2003WO-GB000618.

XX 13-FEB-2002; 2002GB-00003420.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll M, Harrop R, Kingsman S;

XX WPI, 2003-663795/62.

XX P-FSDB; ADB97455.

XX

PT New Major Histocompatibility Complex class II peptide epitope of 5T4,
 PT useful for manufacturing a medicament for diagnosing, preventing and/or
 PT treating a disease, e.g. cancer.

XX Disclosure; Page 49; 63pp; English.

XX The invention relates to a Major Histocompatibility Complex (MHC) class
 CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
 CC clone has a cytostatic activity, as it is useful in manufacturing a
 CC medicament for preventing and/or treating a disease, particularly cancer.
 CC The methods are useful for detecting T-cells capable of specifically
 CC recognising a peptide epitope in conjunction with an MHC molecule, for
 CC diagnosing or monitoring the progression of a cancerous disease, or for
 CC detecting the presence of a peptide or nucleic acid using an agent. The
 CC MHC class II peptide epitope of the invention can be used in gene therapy
 CC or as part of a vaccine. This polynucleotide sequence represents the DNA
 CC coding for the feline 5T4 protein.

SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 22.2 Length: 1260
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ADB97452 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

|||||

Db 1048 TCCCTGCAGACTCTTATGCTTCTTCTA 1074

RESULT 15

AAA27058

ID AAA27058 standard; DNA; 1263 BP.

XX AC AAA27058;

XX 22-AUG-2000 (first entry)

XX Human 5T4 tumour-associated antigen gene.

XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;

XX immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;

XX ds.

XX Homo sapiens.

XX WO200029428-A2.

XX 25-MAY-2000.

XX 18-NOV-1999; 99WO-GB003859.

XX 18-NOV-1998; 98GB-00025303.

XX 27-JAN-1999; 99GB-00001739.

XX 30-JUL-1999; 99GB-00017995.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll MW, Myers KA;

XX WPI; 2000-387735/33.

XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte

XX response useful in vaccinating against and in treating tumors.

XX Example 2; Page 78; 79pp; English.

XX The present sequence encodes the human 5T4 tumour-associated antigen

XX (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in

XX

CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The 5T4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject
 XX
 SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores: 22.2 Length: 1263
 Pred. No.: 43.00 Matches: 9
 Score: 100.0% Conservatives: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 3 Gaps: 0
 DB: 0

US-10-774-176-7 (1-9) x AAA27058 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 1051 TCCCTGCACACCTCTTATGCTTCCTG 1077

RESULT 16
 AAF89736
 ID AAF89736 standard; DNA; 1263 BP.

AC AAF89736;
 XX
 XX 23-JUL-2001 (first entry)
 DT
 DE Nucleotide sequence of canine 5T4 protein.

XX Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
 KW hypersensitivity; autoimmune disease; central nervous system disorder;
 KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
 KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
 KW Helicobacter-related disease; immune disorder; ss.
 XX
 OS Canis sp.

XX Key Location/Qualifiers
 FH 1..1263
 FT CDS /*tag= a
 FT /product= "5T4"
 FT

XX WO200136486-A2.

XX 25-MAY-2001.

XX 13-NOV-2000; 2000WO-GB004317.

XX 18-NOV-1999; 99WO-GB003859.

XX 15-FEB-2000; 2000GB-00003527.

XX 02-MAR-2000; 2000GB-00005071.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;

XX Myers KA;

XX WPI; 2001-343805/36.

XX P-PSDB; AAB83839.

XX Use of single chain antibody capable of recognizing a disease associated

XX molecule for manufacturing a medicament for preventing and/or treating a

XX disease condition associated with disease associated molecule.

XX Disclosure; Fig 26; 118pp; English.

CC The specification describes the use of a single chain antibody (ScFv),
 CC which is capable of recognizing a disease associated molecule in the
 CC manufacture of a medicament for the prevention and treatment of a disease
 CC condition. The ScFv antibody is useful in the manufacture of a
 CC medicament, for affecting a disease in vivo, for preparing a
 CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
 CC treatment of a disease. The ScFv antibody is also useful for treating
 CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
 CC diseases, cancers, central nervous system disorders including Parkinson's
 CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
 CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
 CC related diseases, and other immune disorders. The present sequence
 CC encodes a 5T4 protein, which is used to produce ScFv of the invention
 XX

SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores: 22.2 Length: 1263
 Pred. No.: 43.00 Matches: 9
 Score: 100.0% Conservatives: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 4 Gaps: 0
 DB: 0

US-10-774-176-7 (1-9) x AAF89736 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 1051 TCCCTGCACACCTCTTATGCTTCCTA 1077

RESULT 17
 ABK87174
 ID ABK87174 standard; cDNA; 1263 BP.

AC ABK87174;

XX 07-OCT-2002 (first entry)

DE cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.

XX Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.

XX Canis sp.

XX Key Location/Qualifiers
 FH 1..1263
 FT CDS /*tag= a
 FT /product= "5T4 protein"
 FT

XX WO200238612-A2.

XX 16-MAY-2002.

XX 13-NOV-2001; 2001WO-GB005004.

XX 13-NOV-2000; 2000WO-GB004317.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Myers K, Drury N, Carroll M;

XX WPI; 2002-557449/59.

XX P-PSDB; AAU98693.

XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.

XX Claim 1; Page 67; 68pp; English.

XX

CC The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes canine 5T4 protein
 XX
 SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 22.2 Length: 1263
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x ABK87174 (1-1263)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 DB 1051 TCCCTGCAGACTTCTTATGCTTCTTA 1077

RESULT 18

AAA27059
 ID AAA27059 standard; DNA; 1281 BP.
 XX
 AC AAA27059;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Mouse 5T4 tumour-associated antigen gene.
 XX
 KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
 KW ds.
 XX
 OS Mus musculus.
 XX
 PN WO200029428-A2.
 XX
 PD 25-MAY-2000.
 XX
 PP 18-NOV-1999; 99WO-GB003859.
 XX
 PR 18-NOV-1998; 98GB-00025303.
 PR 27-JAN-1999; 99GB-00001739.
 PR 30-JUL-1999; 99GB-00017995.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Carroll MW, Myers KA;
 XX
 PS WPI; 2000-387735/33.
 XX
 PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.
 XX
 PS Example 2; Page 78; 79pp; English.
 CC
 CC The present sequence encodes the mouse 5T4 tumour-associated antigen
 CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in

CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The 5T4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject. The
 CC present sequence appears in GenBank at accession number AJ012160
 XX
 SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 22.6 Length: 1281
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0

US-10-774-176-7 (1-9) x AAA27059 (1-1281)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 DB 1069 TCCCTGCAGACTTCTTATGCTTCTTA 1095

RESULT 19

AAAD56199
 ID AAAD56199 standard; DNA; 1331 BP.
 XX
 AC AAAD56199;
 XX
 DT 07-AUG-2003 (first entry)
 XX
 DE Human LRRCAPS related DNA #6.
 XX
 KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003035831-A2.
 XX
 PD 01-MAY-2003.
 XX
 PF 21-OCT-2002; 2002WO-US033540.
 XX
 PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX
 PS WPI; 2003-421410/39.
 XX
 PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 PS Disclosure; Page 75-76; 99pp; English.
 XX
 CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a

CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 23.6 Length: 1331
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x AAD56199 (1-1331)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
DB 1081 TCCCTGCAACCTCTTATGCTTCCTG 1107

RESULT 20

ADJ56299
ID ADJ56299 standard; cDNA; 2020 BP.

XX
AC ADJ56299;

XX
DT 06-MAY-2004 (first entry)

XX
DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.

XX
KW human; differential expression; transactivator; proto-oncogene;

XX
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
MYCN activated cell.

XX
OS Homo sapiens.

XX
PN US2003119009-A1.

XX
PD 26-JUN-2003.

XX
PF 25-FEB-2002; 2002US-00084817.

XX
PR 23-FEB-2001; 2001US-0270784P.

XX
PA (STUA/) STUART S G.

XX
PA (NUCH/) NUCHTERN J G.

XX
PA (PLON/) PLON S E.

XX
PA (SHOH/) SHOHET J M.

XX
PI Stuart SG, Nuchtern JG, Pion SE, Shohet JM;

XX
DR WPI; 2003-635698/60.

XX
PT New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.

XX
PS Claim 1; SEQ ID NO 105; 27pp; English.

XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in

CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.

XX
SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 37.7 Length: 2020
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ADJ56299 (1-2020)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
DB 1121 TCCCTGCAACCTCTTATGCTTCCTG 1147

RESULT 21

ADJ51052

ID ACC51052 standard; cDNA; 2053 BP.

XX
AC ACC51052;

XX
DT 12-JUN-2003 (first entry)

XX
DE Human bladder cancer associated cDNA sequence SEQ ID NO:192.

XX
KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.

XX
OS Homo sapiens.

XX
PN WO2003003906-A2.

XX
PD 16-JAN-2003.

XX
PF 03-JUL-2002; 2002WO-US021338.

XX
PR 03-JUL-2001; 2001US-0302814P.

XX
PR 03-AUG-2001; 2001US-0310099P.

XX
PR 08-NOV-2001; 2001US-0343705P.

XX
PR 13-NOV-2001; 2001US-0350666P.

XX
PR 12-APR-2002; 2002US-0372246P.

XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX
PI Mack DH, Aziz N;

XX
DR WPI; 2003-201532/19.

XX
DR P-PSDB; ABR48236.

XX
PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.

XX
PS Claim 6; Page 296; 307pp; English.

XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridises to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine

CC and/or antisense/inhibition applications

XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	38.4	Length:	2053
Score:	43.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0

US-10-774-176-7 (1-9) x ACCS1052 (1-2053)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1135 TCCCTGCAACCTCTTATGCTTCTG 1161

RESULT 22

ABX76332

ID ABX76332 standard; DNA; 2053 BP.

AC ABX76332;

XX 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #196.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;

KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

OS WO20286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI, 2003-093161/08.

DR P-ESDB; ABUS6603.

XX Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased

XX expression in lung cancer.

XX Claim 22; Page 335; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated

CC transcript in a cell from a patient, comprising contacting a biological

CC sample from the patient with a polynucleotide that selectively hybridises

CC to a sequence that is at least 80 % identical to a gene that exhibits

CC increased or decreased expression in lung cancer samples. Lung cancer-

CC associated polynucleotides and polypeptides are used for identifying a

CC compound that modulates a lung cancer-associated polypeptide, for

CC inhibiting proliferation of a lung cancer-associated cell to treat lung

CC cancer in a patient and for treating a mammal having lung cancer by

CC administering a modulatory compound identified. The methods are useful

CC for treating lung cancer, such as small cell lung cancer, non-small cell

CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,

CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,

CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and

CC bronchiectasis. The genes, polynucleotides and polypeptides are useful

CC for diagnostic purposes and as targets for screening for therapeutic

CC compounds that modulate lung cancer, such as antibodies. Sequences

CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the

XX invention

XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	38.4	Length:	2053
Score:	43.00 <td>Matches:</td> <td>9 </td>	Matches:	9
Percent Similarity:	100.0% <td>Conservative:</td> <td>0 </td>	Conservative:	0
Best Local Similarity:	100.0% <td>Mismatches:</td> <td>0 </td>	Mismatches:	0
Query Match:	100.0% <td>Indels:</td> <td>0 </td>	Indels:	0
DB:	8 <td>Gaps:</td> <td>0 </td>	Gaps:	0

US-10-774-176-7 (1-9) x ABX76332 (1-2053)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1135 TCCCTGCAACCTCTTATGCTTCTG 1161

RESULT 23

AAD56197

ID AAD56197 standard; DNA; 2053 BP.

XX AAD56197;

AC 07-AUG-2003 (first entry)

XX Human LRRCAPS DNA #11.

XX Human; p53 pathway; Leucine rich repeat capricious related protein;

KW LRRCAPS; cancer; gene therapy; ds.

XX Homo sapiens.

XX WO2003035831-A2.

XX 01-MAY-2003.

XX 21-OCT-2002; 2002WO-US033540.

XX 22-OCT-2001; 2001US-0338733P.

PR 15-FEB-2002; 2002US-0357600P.

PR 01-MAR-2002; 2002US-0361196P.

XX (EXEL-) EXELIXIS INC.

XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;

PI Francis-Lang H, Friedman L;

XX WPI, 2003-421410/39.

XX Identifying a candidate p53 pathway-modulating agent for treating cancer

PT comprises contacting an assay system comprising a purified leucine rich

PT repeat, capricious related polypeptide or nucleic acid with a test agent.

XX Example 5; Page 73-74; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway

CC modulating agent. The method involves contacting an assay system

CC comprising a purified Leucine rich repeat, capricious related (LRRCAPS)

CC polypeptide or nucleic acid or its fragment with a test agent and

CC detecting a test agent-biased activity, where a difference between the

CC test agent-biased activity and the reference activity identifies the test

CC agent as a candidate p53 pathway modulating agent. The method is useful

CC for identifying a candidate p53 pathway-modulating agent for preparing a

CC composition for diagnosing or treating cancer. The invention is useful in

CC gene therapy. The present sequence is human LRRCAPS DNA
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 38.4 Length: 2053
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x AAD56197 (1-2053)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 DB 1135 TCCTGCAAAACCTCTTATGTCCTCG 1161

RESULT 24

AAD56200
 ID AAD56200 standard; DNA; 2053 BP.

XX AC AAD56200;

DT 07-AUG-2003 (first entry)

XX DE Human LRRCAPS DNA #12.

XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.

XX OS Homo sapiens.

XX FN WO2003035831-A2.

XX PD 01-MAY-2003.

XX PF 21-OCT-2002; 2002WO-US033540.

XX PR 22-OCT-2001; 2001US-0338733P.

XX PR 15-FEB-2002; 2002US-0357600P.

XX PR 01-MAR-2002; 2002US-0361196P.

XX PA (EXEL-) EXELIXIS INC.

XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioulin MN, Li D;
 PI Francis-Lang H, Friedman L;

XX DR WPI; 2003-421410/39.

XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.

XX PS Disclosure; Page 76-77; 99pp; English.

XX CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 38.4 Length: 2053
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x AAD56200 (1-2053)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 DB 1135 TCCTGCAAAACCTCTTATGTCCTCG 1161

RESULT 25

ADN38721

ID ADN38721 standard; cDNA; 2053 BP.

XX AC ADN38721;

XX DT 17-JUN-2004 (first entry)

XX KW Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.

XX KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnerary; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX PN WO2003042661-A2.

XX XX 22-MAY-2003.

XX PF 13-NOV-2002; 2002WO-US036810.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 21-NOV-2001; 2001US-0332464P.

XX PR 29-NOV-2001; 2001US-0334393P.

XX PR 03-DEC-2001; 2001US-0335394P.

XX PR 14-DEC-2001; 2001US-0340376P.

XX PR 08-JAN-2002; 2002US-0347211P.

XX PR 10-JAN-2002; 2002US-0347349P.

XX PR 08-FEB-2002; 2002US-0355250P.

XX PR 13-FEB-2002; 2002US-0356714P.

XX PR 20-FEB-2002; 2002US-0359077P.

XX PR 29-MAR-2002; 2002US-0368809P.

XX PR 04-APR-2002; 2002US-0370110P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PR 05-JUN-2002; 2002US-0386614P.

XX PR 16-JUL-2002; 2002US-0396839P.

XX PR 22-JUL-2002; 2002US-0397775P.

XX PR 23-JUL-2002; 2002US-0397845P.

XX PR 09-SEP-2002; 2002US-0409450P.

XX XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX PA Afar D, Aziz N, Gineburg WM, Gish KC, Glynne R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX DR WPI; 2003-468649/44.

XX DR P-PSDB; ADN38722.

XX PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.

XX PS Claim 8; SEQ ID NO 39; 1385pp; English.

XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a

CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 38.4 Length: 2053
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-7 (1-9) x ADN38721 (1-2053)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
|||||
Db 1135 TCCTGCAACCTCTTATGTCCTCTG 1161

RESULT 26

ADL06473
ID ADL06473 standard; cDNA; 2053 BP.

XX AC

XX ADL06473;

XX 20-MAY-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #53.
XX Human; tumour-associated antigenic target; TAT; cell death; tumour;
XX cancer; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004016225-A2.

XX 26-FEB-2004.

XX 19-AUG-2003; 2003WO-US025892.

XX 19-AUG-2002; 2002US-0404809P.

XX 21-AUG-2002; 2002US-0405645P.

XX 23-SEP-2002; 2002US-0413192P.

XX 15-OCT-2002; 2002US-0419008P.

XX 15-NOV-2002; 2002US-0426847P.

XX 02-JUL-2003; 2003US-0484959P.

XX (GETH) GENENTECH INC.

XX Desauvage FJ, Prantz G, Hillan KJ, Polakis P, Polson A, Smith V;
XX Spencer SD, Wu TD, Zhang Z;

XX WPI; 2004-257144/24.

XX P-PSDB; ADL06552.

XX New antibody that binds to a tumor-associated antigenic target (TAT)
XX polypeptide, useful for preparing a composition for diagnosing or
XX treating cancer.

XX Claim 1; SEQ ID NO 53; 319pp; English.

XX The present invention relates to the isolation of human tumour-associated

CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also
CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is
CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a
CC humanised antibody. It is conjugated to a growth inhibitory agent. It is
CC produced in bacteria or in CHO cells and induces death of a cell to which
CC it binds. The antibody is useful for preparing a composition for
CC diagnosing or treating tumours and cancer. The present sequence
CC represents a human TAT cDNA sequence of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 38.4 Length: 2053
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-7 (1-9) x ADL06473 (1-2053)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
|||||
Db 1135 TCCTGCAACCTCTTATGTCCTCTG 1161

RESULT 27

ADN03961

ID ADN03961 standard; cDNA; 2053 BP.

XX AC

XX ADN03961;

XX 01-JUL-2004 (first entry)

XX DE Antipsoriatic cDNA sequence #180.

XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

XX WO2004028479-A2.

XX 08-APR-2004.

XX 25-SEP-2003; 2003WO-US030907.

XX 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.

XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX Wu TD;

XX WPI; 2004-305105/28.

XX P-PSDB; ADN03962.

XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.

XX Claim 1; SEQ ID NO 355; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.

SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 38.4 Length: 2053
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-7 (1-9) x ADN03961 (1-2053)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 Db 1135 TCCCTGCAACCTCTATGTCCTCTG 1161

RESULT 28

ADR25444

ID ADR25444 standard; DNA; 2053 BP.

XX AC ADR25444;

XX DT 21-OCT-2004 (first entry)

XX DE Breast cancer prognosis marker #1305.

XX KW ds; breast cancer; prognosis; gene expression; diagnosis.

XX OS Homo sapiens.

XX PN WO2004065545-A2.

XX PD 05-AUG-2004.

XX PF 15-JAN-2004; 2004WO-US001100.

XX PR 15-JAN-2003; 2003US-00342887.

XX PA (ROSE-) ROSETTA INPHARMATICS LLC.

XX PA (NECA-) NETHERLANDS CANCER INST.

XX PI Van't Veer LJ, He Y;

XX DR WPI; 2004-593473/57.

XX PT Classifying a breast cancer patient according to prognosis comprises

XX PT determining the similarity between the level of expression of each of

XX PT five genes in a cell sample taken from patient, to control levels.

XX PS Disclosure; SEQ ID NO 1305; 226pp; English.

XX CC The invention relates to a method of classifying a breast cancer patient

XX CC according to prognosis by determining the similarity between the level of

XX CC expression of each of five genes for which markers are listed in the

XX CC specification, in a cell sample taken from the breast cancer patient, to

XX CC control levels of expression for each respective five genes to obtain a

XX CC patient similarity value. The methods are useful for classifying a breast

XX CC cancer patient according to prognosis. Kits and computer program products

XX CC are useful for data analysis using the diagnostic, prognostic and

XX CC statistical methods of the invention. This sequence corresponds to a

XX CC marker used in the method of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 38.4 Length: 2053

Score: 43.00 Matches: 9

Percent Similarity: 100.0% Mismatches: 0

Best Local Similarity: 100.0% Indels: 0

Query Match: 100.0% Gaps: 0

DB: 13

US-10-774-176-7 (1-9) x ADR25444 (1-2053)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1135 TCCCTGCAACCTCTATGTCCTCTG 1161

RESULT 29

ACN38510
 ID ACN38510 standard; cDNA; 2053 BP.
 XX AC ACN38510;
 XX DT 18-NOV-2004 (first entry)
 XX DE Tumour-associated antigenic target (TAT) cDNA DNA103471, SEQ ID NO:2070.
 XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 XX KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 XX KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 XX KW central nervous system cancer; bladder cancer; pancreatic cancer;
 XX KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 XX KW chromosome identification; chromosome mapping; gene mapping;
 XX KW gene therapy; cytostatic; gene; ss.

XX OS Homo sapiens.

XX PN WO2004030615-A2.

XX PD 15-APR-2004.

XX PF 29-SEP-2003; 2003WO-US028547.

XX PR 02-OCT-2002; 2002US-0414971P.

XX PA (GETH) GENENTECH INC.

XX PI Wu TD, Zhang Z, Zhou Y;

XX DR WPI; 2004-347921/32.

XX DR P-PSDB; ABM80804.

XX PT New tumor-associated antigenic target polypeptides and nucleic acids,
 XX PT useful in preparing a medicament for treating or detecting a
 XX PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 XX PT prostate cancer or tumor.

XX Claim 1; SEQ ID NO 2070; 7273pp; English.

XX CC The invention relates to human tumour-associated antigenic target (TAT)
 XX CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 XX CC overexpressed in cancer tissues compared to normal tissues, and may thus
 XX CC serve as effective targets for the diagnosis and treatment of cancer in
 XX CC mammals. The invention also relates to nucleic acid and polypeptide
 XX CC sequences at least 80% identical to the TAT nucleic acids and
 XX CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 XX CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 XX CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 XX CC TAT polypeptide; and methods and compositions for the treatment or
 XX CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 XX CC antibodies, antagonists, binding molecules and compositions are useful
 XX CC for diagnosing or treating a cell proliferative disorder associated with
 XX CC increased TAT expression, particularly cancers such as breast cancer,
 XX CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 XX CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 XX CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 XX CC used as hybridisation probes, in chromosome and gene mapping, in
 XX CC chromosome identification and in gene therapy. The present sequence
 XX CC represents a TAT nucleic acid of the invention

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 38.4 Length: 2053

Score: 43.00 Matches: 9

Percent Similarity: 100.0% Mismatches: 0

Best Local Similarity: 100.0% Indels: 0

Query Match: 100.0% Gaps: 0

DB: 13

US-10-774-176-7 (1-9) x ACN38510 (1-2053)


```
Qy      1 SerLeuGlnThrSerTy-ValPheLeu 9
      |||||
Db      1135 TCCCTGCAAACTCTTATGCTTCCTG 1161

RESULT 30
ADV35098
ID      ADV35098 standard; cDNA; 2053 BP.
XX
AC      ADV35098;
XX
DT      10-FEB-2005 (first entry)
XX
DE      Human cDNA of an exemplary efficacy gene for BAD SeqID174.
XX
KW      human; ss; multi-parameter high throughput screening; MPHTS;
KW      disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
KW      bipolar affective disorder; BAD; autism; Parkinson's;
KW      Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.
XX
OS      Homo sapiens.
XX
FN      US2003096264-A1.
XX
PD      22-MAY-2003.
XX
PF      18-JUN-2002; 2002US-00175523.
XX
PR      07-SEP-2001; 2001US-0299151P.
PR      07-SEP-2001; 2001US-0317828P.
PR      25-SEP-2001; 2001US-0325150P.
PR      14-NOV-2001; 2001US-0333047P.
PR      18-JAN-2002; 2002US-0349936P.
PR      04-MAR-2002; 2002US-0361834P.
XX
PA      (PSYC-) PSYCHIATRIC GENOMICS INC.
XX
PI      Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
PI      Palfreyman M, Rajan P;
XX
WPI; 2004-118903/12.
XX
PT      Identifying a compound that can treat disease or disorders, such as, a
PT      neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
PT      determining the expression of one or more efficacy genes in a cell
PT      contacted with the test compound.
XX
PS      Example 6; SEQ ID NO 174; 39pp; English.
XX
CC      This invention relates to a novel screening method identified as a multi-
CC      parameter high throughput screening (MPHTS) assay. Specifically, it
CC      refers to an assay that utilizes the disease signature of a plurality of
CC      specific genes associated with a particular disease, and identifies
CC      differential expression between those cells taken from individuals
CC      affected by that disease and those that are not affected. The present
CC      invention then describes the screening of candidate pharmaceutical
CC      compounds to identify those that have a potential therapeutic benefit for
CC      the treatment of neuropsychiatric and neurodegenerative disorders
CC      including schizophrenia, bipolar affective disorder (BAD) and autism, as
CC      well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
CC      of this invention exhibit various activities including neuroleptic,
CC      nootropic, antimanic and antidepressant. Furthermore, the screening
CC      method used in MPHTS will be automated, such that a large number of test
CC      compounds may be rapidly screened with a minimal amount of labour and
CC      effort. This polynucleotide is a human cDNA sequence of a gene that is
CC      differentially expressed in the presence of a therapeutic compound and
CC      represents an exemplary efficacy gene for bipolar affective disorder,
CC      given in an exemplification of the invention.
XX
SQ      Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:      38.4      Length:      2053
Pred. No.:
```

```
Score:      43.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match: 100.0%      Indels: 0
DB: 13      Gaps: 0

US-10-774-176-7 (1-9) x ADV35098 (1-2053)

Qy      1 SerLeuGlnThrSerTy-ValPheLeu 9
      |||||
Db      1135 TCCCTGCAAACTCTTATGCTTCCTG 1161

RESULT 31
AAS87175
ID      AAS87175 standard; cDNA; 2338 BP.
XX
AC      AAS87175;
XX
DT      13-FEB-2002 (first entry)
XX
DE      DNA encoding novel human diagnostic protein #22979.
XX
KW      Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW      food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS      Homo sapiens.
XX
PN      WO200175067-A2.
XX
PD      11-OCT-2001.
XX
PF      30-MAR-2001; 2001WO-US008631.
PR      31-MAR-2000; 2000US-00540217.
PR      23-AUG-2000; 2000US-00649167.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Drmanac RT, Liu C, Tang YT;
XX
WPI; 2001-639362/73.
DR      P-PSDB; ABG22988.
XX
PT      New isolated polynucleotide and encoded polypeptides, useful in
PT      diagnostics, forensics, gene mapping, identification of mutations
PT      responsible for genetic disorders or other traits and to assess
PT      biodiversity.
XX
PS      Claim 1; SEQ ID NO 22979; 103pp; English.
XX
CC      The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC      sequences. (I) is useful as hybridisation probes, polymerase chain
CC      reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC      and in recombinant production of (II). The polynucleotides are also used
CC      in diagnostics as expressed sequence tags for identifying expressed
CC      genes. (I) is useful in gene therapy techniques to restore normal
CC      activity of (II) or to treat disease states involving (II). (II) is
CC      useful for generating antibodies against it, detecting or quantitating a
CC      polypeptide in tissue, as molecular weight markers and as a food
CC      supplement. (II) and its binding partners are useful in medical imaging
CC      of sites expressing (II). (I) and (II) are useful for treating disorders
CC      involving aberrant protein expression or biological activity. The
CC      polypeptide and polynucleotide sequences have application in
CC      diagnostics, forensics, gene mapping, identification of mutations
CC      responsible for genetic disorders or other traits to assess biodiversity
CC      and to produce other types of data and products dependent on DNA and
CC      amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC      coding sequences of the invention. Note: The sequence data for this
CC      patent did not appear in the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;
```

Alignment Scores:
 Pred. No.: 44.5 Length: 2338
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0

US-10-774-176-7 (1-9) x AAS87175 (1-2338)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 DB 1392 TCCCTGCAAACTCTTATGCTTCCTG 1418

RESULT 32

AAK94253

ID AAK94253 standard; cDNA; 2359 BP.

XX AC AAK94253;

DT 06-NOV-2001 (first entry)

DE Human full-length cDNA, SEQ ID NO: 2864.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX OS Homo sapiens.

PN EP1130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-00114089.

XX PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

XX PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

DR P-PSDB; AAM93333.

PT 830 Primers useful for synthesizing full length cDNA clones and their use
 in genetic manipulation.

XX PS Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.

XX CC The invention relates to primers for synthesizing full length cDNA
 clones. 830 cDNA molecules encoding a human protein have been isolated
 and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 been determined. Primers for synthesizing the full length cDNA are useful
 for clarifying the function of the protein encoded by the cDNA. The full
 length clones were obtained by construction of full length enriched cDNA
 libraries that were synthesised by the oligo-capping method. The primers
 enable the production of the full length cDNA easily without any special
 methods. The present sequence is a full length human cDNA of the
 invention. Note: The sequence data for this patent did not form part of
 the printed specification, but was obtained in CD-ROM format directly
 from BPO

XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 45 Length: 2359
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0

DB: 4 Gaps: 0

US-10-774-176-7 (1-9) x AAK94253 (1-2359)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 DB 1474 TCCCTGCAAACTCTTATGCTTCCTG 1500

RESULT 33

ADL30831

ID ADL30831 standard; cDNA; 2359 BP.

XX AC ADL30831;

DT 20-MAY-2004 (first entry)

XX DE Full length human cDNA clone SeqID 2864.

XX KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.

XX OS Homo sapiens.

XX PN EP1396543-A2.

XX PD 10-MAR-2004.

XX PF 07-JUL-2000; 2003EP-00025638.

XX PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183865.

PR 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

DR P-PSDB; ADL30832.

PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 length human cDNAs.

XX PS Example 1; SEQ ID NO 2864; 1340pp; English.

XX CC This invention relates to a novel primers useful for synthesising full
 length cDNA molecules that encode human proteins. Specifically, it refers
 to secretory or membrane proteins that are potential therapeutic agents/
 target molecules in the field of medicine, and in particular genes
 encoding proteins that are associated with signal transduction.
 CC glycoproteins and transcription. The present invention describes a method
 for efficiently cloning a full length human cDNA from both the 5' and 3'
 ends using the oligo-capping method. This polynucleotide sequence is a
 full length human cDNA clone of the invention.

XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 45 Length: 2359
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-7 (1-9) x ADL30831 (1-2359)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||

DB 1474 TCCCTGCAAACTCTTATGCTTCCTG 1500

RESULT 34
AAK94254
ID AAK94254 standard; cDNA; 2361 BP.
XX
AC AAK94254;
AC AC
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2866.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD PD
XX
PF 05-SEP-2001.
XX
PR 07-JUL-2000; 2000BP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
XX
PR 11-JAN-2000; 2000JP-00118774.
XX
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
DR WPI; 2001-524255/58.
DR P-PSDB; AAM93334.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their use
in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
clones. 830 cDNA molecules encoding a human protein have been isolated
and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
been determined. Primers for synthesising the full length cDNA are useful
for clarifying the function of the protein encoded by the cDNA. The full
length clones were obtained by construction of full length enriched cDNA
libraries that were synthesised by the oligo-capping method. The primers
enable the production of the full length cDNA easily without any special
methods. The present sequence is a full length human cDNA of the
invention. Note: The sequence data for this patent did not form part of
the printed specification, but was obtained in CD-ROM format directly
from EPO
XX
SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 45 Length: 2361
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-7 (1-9) x AAK94254 (1-2361)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1476 TCCTCGAACCCTATTATGTCTTCGG 1502

RESULT 35
ADI26162
ID ADI26162 standard; cDNA; 2361 BP.
XX
AC ADI26162;
XX
DT 22-APR-2004 (first entry)

```

QY      1 SerLeuGlnThrSerTyrValPheLeu 9
Db      1476 TCCCTGCAAACTCTTATGTCCTCTG 1502

RESULT 36
ADL30833
ID      ADL30833 standard; cDNA; 2361 BP.
XX
AC      ADL30833;
XX
DT      20-MAY-2004 (first entry)
XX
DE      Full length human cDNA clone SeqID 2866.
XX
KW      human; medicine; signal transduction; glycoprotein; transcription;
KW      oligo-capping method; ss; gene.
XX
OS      Homo sapiens.
XX
PN      EP1396543-A2.
XX
PD      10-MAR-2004.
XX
PF      07-JUL-2000; 2003EP-00025638.
XX
PR      08-JUL-1999; 99JP-00194486.
PR      11-JAN-2000; 2000JP-00118774.
PR      02-MAY-2000; 2000JP-00183865.
PR      07-JUL-2000; 2000EP-00114089.
XX
PA      (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI      Ota T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;
PI      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
WPI; 2004-204755/20.
DR      P-PSDB; ADL30834.
XX
PT      New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT      length human cDNAs.
XX
PS      Example 1; SEQ ID NO 2866; 1340pp; English.
XX
CC      This invention relates to a novel primers useful for synthesizing full
CC      length cDNA molecules that encode human proteins. Specifically, it refers
CC      to secretory or membrane proteins that are potential therapeutic agents/
CC      target molecules in the field of medicine, and in particular genes
CC      encoding proteins that are associated with signal transduction,
CC      glycoproteins and transcription. The present invention describes a method
CC      for efficiently cloning a full length human cDNA from both the 5' and 3'
CC      ends using the oligo-capping method. This polynucleotide sequence is a
CC      full length human cDNA clone of the invention.
XX
SQ      Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 45 Length: 2361
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-7 (1-9) x ADL30833 (1-2361)

QY      1 SerLeuGlnThrSerTyrValPheLeu 9
Db      1476 TCCCTGCAAACTCTTATGTCCTCTG 1502

RESULT 37
ADL26160
ID      ADL26160 standard; cDNA; 2557 BP.

```

```

XX      ADL26160;
XX      22-APR-2004 (first entry)
XX      Human cDNA encoding protein that promotes STAT6 activation #63.
XX
DE      ss; gene; human; signal transducer and activator of transcription 6;
DE      STAT6; immunogen; STAT6 activation; allergy; inflammation;
KW      autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
KW      Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
KW      systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
KW      ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
XX
OS      Homo sapiens.
XX
PN      WO2003104277-A2.
XX
PD      18-DEC-2003.
XX
PF      05-JUN-2003; 2003WO-JP007123.
XX
PR      05-JUN-2002; 2002JP-00164257.
PR      06-JUN-2002; 2002US-0385912P.
PR      26-DEC-2002; 2002JP-00377326.
PR      27-DEC-2002; 2002US-0436467P.
PR      15-MAY-2003; 2003JP-00137505.
PR      16-MAY-2003; 2003US-0470836P.
XX
PA      (ASAH ) ASAH KASEI KK.
XX
PI      Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
XX
WPI; 2004-122214/12.
DR      P-PSDB; ADL26161.
XX
PT      New signal transducer and activator of transcription 6 activation
PT      promoting purified protein, for diagnosing and treating disease
PT      associated with activation/inhibition of transcription factor e.g.
PT      diabetes and cancer.
XX
PS      Claim 4; SEQ ID NO 125; 1368pp; English.
XX
CC      The invention relates to a purified protein promoting signal transducer
CC      and activator of transcription 6 activation (STAT6). The protein is
CC      useful for the producing an antibody, which involves administering the
CC      protein or its epitope-bearing fragments to a non-human animal as an
CC      antigen. The nucleic acid is useful for diagnosing a disease or
CC      susceptibility to a disease related to expression or activity of the
CC      protein. A transformant expressing the protein is useful for screening
CC      compounds which inhibit or promote STAT6 activation. A transformant
CC      expressing the protein is useful for producing a pharmaceutical
CC      composition. Compositions, antibodies and antisense molecules are useful
CC      for the treating a disease associated with STAT6 activation such as
CC      allergic diseases, inflammation, autoimmune diseases, diabetes,
CC      hyperlipidaemia, infections disease and cancers. Compositions are useful
CC      for treating disease associated with STAT6 activation and/or prevention
CC      of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC      arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC      allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC      viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC      activity. The protein or nucleic acid is effectively useful for screening
CC      compounds for treating and preventing disease associated with excessive
CC      activation or inhibition of STAT6. The present sequence represents a
CC      human cDNA encoding a protein which promotes STAT6 activation.
XX
SQ      Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 49.2 Length: 2557
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

```


CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
SQ Sequence 14952 BP; 3174 A; 3899 C; 4409 G; 3470 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.59e+03 Length: 14952
Score: 39.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.7% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ACF69283 (1-14952)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
||||:|||||
Db 8207 TCTATCCAGACCAGTTATGTTATTCATC 8233

RESULT 40
ACF65386_6
Continuation (7 of 7) of ACF65386 from base 600001 (Photorhabdus luminescens nucleotide
WP Sequence split into 7 fragments LOCUS ACF65386 Accession ACF65386
WP Fragment Name Begin End
WP ACF65386_0 1 110000
WP ACF65386_1 100001 210000
WP ACF65386_2 200001 310000
WP ACF65386_3 300001 410000
WP ACF65386_4 400001 510000
WP ACF65386_5 500001 610000
WP ACF65386_6 600001 700779

Alignment Scores:
Pred. No.: 2.22e+04 Length: 100779
Score: 39.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.7% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ACF65386_6 (1-100779)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
||||:|||||
Db 80885 TCTATCCAGACCAGTTATGTTATTCATC 80911

RESULT 41
ACF67367_22/c
Continuation (23 of 57) of ACF67367 from base 2200001 (Photorhabdus luminescens nucleotide
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367
WP Fragment Name Begin End
WP ACF67367_00 1 110000
WP ACF67367_01 100001 210000
WP ACF67367_02 200001 310000
WP ACF67367_03 300001 410000
WP ACF67367_04 400001 510000
WP ACF67367_05 500001 610000
WP ACF67367_06 600001 710000
WP ACF67367_07 700001 810000
WP ACF67367_08 800001 910000
WP ACF67367_09 900001 1010000
WP ACF67367_10 1000001 1110000
WP ACF67367_11 1100001 1210000
WP ACF67367_12 1200001 1310000
WP ACF67367_13 1300001 1410000
WP ACF67367_14 1400001 1510000
WP ACF67367_15 1500001 1610000
WP ACF67367_16 1600001 1710000

WP ACF67367_17 1700001 1810000
WP ACF67367_18 1800001 1910000
WP ACF67367_19 1900001 2010000
WP ACF67367_20 2000001 2110000
WP ACF67367_21 2100001 2210000
WP ACF67367_22 2200001 2310000
WP ACF67367_23 2300001 2410000
WP ACF67367_24 2400001 2510000
WP ACF67367_25 2500001 2610000
WP ACF67367_26 2600001 2710000
WP ACF67367_27 2700001 2810000
WP ACF67367_28 2800001 2910000
WP ACF67367_29 2900001 3010000
WP ACF67367_30 3000001 3110000
WP ACF67367_31 3100001 3210000
WP ACF67367_32 3200001 3310000
WP ACF67367_33 3300001 3410000
WP ACF67367_34 3400001 3510000
WP ACF67367_35 3500001 3610000
WP ACF67367_36 3600001 3710000
WP ACF67367_37 3700001 3810000
WP ACF67367_38 3800001 3910000
WP ACF67367_39 3900001 4010000
WP ACF67367_40 4000001 4110000
WP ACF67367_41 4100001 4210000
WP ACF67367_42 4200001 4310000
WP ACF67367_43 4300001 4410000
WP ACF67367_44 4400001 4510000
WP ACF67367_45 4500001 4610000
WP ACF67367_46 4600001 4710000
WP ACF67367_47 4700001 4810000
WP ACF67367_48 4800001 4910000
WP ACF67367_49 4900001 5010000
WP ACF67367_50 5000001 5110000
WP ACF67367_51 5100001 5210000
WP ACF67367_52 5200001 5310000
WP ACF67367_53 5300001 5410000
WP ACF67367_54 5400001 5510000
WP ACF67367_55 5500001 5610000
WP ACF67367_56 5600001 5648894

Alignment Scores:
Pred. No.: 2.46e+04 Length: 110000
Score: 39.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.7% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ACF67367_22 (1-110000)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
||||:|||||
Db 17741 TCTATCCAGACCAGTTATGTTATTCATC 17715

RESULT 42
AD212550/c
ID AD212550 standard; DNA; 117730 BP.
XX
AC AD212550;
XX
DT 16-JUN-2005 (first entry)
XX
DE Human cancer-associated genomic DNA #8.
XX
KW Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm;
XX cytostatic; gene; ds.
OS Homo sapiens.
XX
PN WO2005031001-A2.
XX
PD 07-APR-2005.

XX 23-SEP-2004; 2004WO-US031617.
 XX 23-SEP-2003; 2003US-00669920.
 XX (CHIR) CHIRON CORP.
 XX Morris DW, Malandro MS;
 XX WPI; 2005-273395/28.
 XX Nucleic acid array useful for detecting cancer associated nucleic acid,
 PT comprises two or more nucleic acid probes.
 XX
 PS Disclosure; SEQ ID NO 70; 198pp; English.
 XX
 CC The invention relates to a nucleic acid array for detecting a cancer
 CC associated (CA) nucleic acid, comprising two or more nucleic acid probes.
 CC The invention also relates to a peptide array comprising two or more
 CC isolated polypeptides encoded by a CA nucleic acid sequence, a compound
 CC that binds to a polypeptide, an isolated antibody or its fragment which
 CC binds to a polypeptide, which is prepared by immunizing a host animal
 CC with a composition comprising the polypeptide or its antigen binding
 CC fragment and collecting cells from the host expressing antibodies against
 CC the antigen or its antigen binding fragment, a composition comprising the
 CC antibody and a carrier, a method of screening for anticancer activity, a
 CC method of detecting a CA nucleic acid, a method of diagnosing cancer, a
 CC method of treating cancer and a method of inhibiting expression of a CA
 CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
 CC nucleic acids. The antibody is useful for detecting the presence or
 CC absence of cancer cells in an individual which involves contacting cells
 CC from the individual with the antibody and detecting a complex of a CA
 CC protein from the cancer cells and the antibody, where the detection of
 CC the complex correlates with the presence of cancer cells in the
 CC individual. The composition is useful for inhibiting growth of cancer
 CC cells in an individual or for delivering a therapeutic agent to cancer
 CC cells in an individual. The invention is also useful for diagnosing
 CC cancer, for treating cancer and for inhibiting expression of a CA gene in
 CC a cell. This sequence represents human cancer-associated genomic DNA of
 CC the invention.
 XX
 SQ Sequence 117730 BP; 32629 A; 25620 C; 25334 G; 34147 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.34e+04 Length: 117730
 Score: 38.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 88.4% Indels: 0
 DB: 14 Gaps: 0
 US-10-774-176-7 (1-9) x ADZ12550 (1-117730)
 Qy 2 LeuGlnThrSerTyrValPheLeu 9
 Db 16982 CTTGAGCTTCATATATTTCTT 16959
 RESULT 43
 ID ADX45743
 ID ADX45743 standard; cDNA; 565 BP.
 XX ADX45743;
 AC ADX45743;
 XX
 XX 21-APR-2005 (first entry)
 DT
 XX Plant full length insert polynucleotide seqid 20483.
 DE
 XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;

KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content; gene; ss.
 XX Unidentified.
 XX US2004034888-A1.
 XX 19-FEB-2004.
 XX 28-APR-2003; 2003US-00425114.
 XX 06-MAY-1999; 99US-00304517.
 PR 05-NOV-2001; 2001US-00985678.
 XX (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOY/) CAO Y.
 XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
 PI WPI; 2004-180133/17.
 XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.
 XX Claim 1; SEQ ID NO 20483; 15pp; English.
 XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This sequence represents a plant full length insert
 CC polynucleotide that can be used in the recombinant DNA construct of the
 CC invention.

XX SQ Sequence 565 BP; 128 A; 153 C; 157 G; 127 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 172 Length: 565
 Score: 37.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 86.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-7 (1-9) x ADX45743 (1-565)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 401 AGCCTGCAACCTTCGTACGTTACGTA 427

RESULT 44
 AAT75704
 ID AAT75704 standard; cDNA; 2848 BP.
 XX
 AC AAT75704;
 XX

DT 16-MAR-1998 (first entry)
 XX Murine leptin receptor splice variant OB-Rb encoding cDNA.
 XX
 DE Murine, leptin receptor; OB-R; obesity; diabetes; high blood pressure;
 KW high cholesterol; body weight; ds.
 KW
 XX Mus musculus.
 OS
 XX WO9726335-A1.
 PN
 XX 24-JUL-1997.
 PD
 XX
 XX 16-JAN-1997; 97WO-US001010.
 XX
 PF 16-JAN-1996; 96US-00586594.
 PR 14-FEB-1996; 96US-00599974.
 PR
 XX (UVRQ) UNIV ROCKEFELLER.
 PA
 XX Friedman JM, Lee G, Proenca R, Ioffe E;
 PI WPI; 1997-385338/35.
 XX
 CC Leptin receptor, OB-R, polypeptide - useful to treat obesity, optionally
 CC in conjunction with treatment for diabetes, high blood pressure and high
 CC cholesterol.
 PT
 XX
 PS Claim 28; Page 98-99; 171pp; English.
 XX
 CC The present sequence encodes a leptin receptor (OB-R) protein, OB-Rb. The
 CC OB-R can be used to treat obesity, optionally in conjunction with a
 CC treatment for diabetes, high blood pressure and high cholesterol, or in
 CC cosmetic compositions for reducing body weight. It may also be used in
 CC agriculture to produce leaner food animals, e.g. beef cattle, swine
 CC poultry, sheep. An antibody specific for OB-R can be used to measure the
 CC presence of OB-R in a sample, optionally in vivo, while the nucleic acid
 CC molecule encoding OB-R can be used to detect defects in the OB-R
 CC polypeptide associated with obese phenotypes, or diagnostically to
 CC measure its encoded RNA and protein in nutritional disorders. The nucleic
 CC acid molecule can be used in gene therapy, or the antisense nucleic acid
 CC molecule can be used to antagonise leptin activity. The nucleic acid
 CC molecule, or the antisense nucleic acid molecule, can be used to treat
 CC weight loss e.g. associated with AIDS, cancer or anorexia nervosa
 XX
 SQ Sequence 2848 BP; 785 A; 596 C; 623 G; 829 T; 0 U; 15 Other;
 PS
 Alignment Scores:
 Pred. No.: 1.07e+03 Length: 2848
 Score: 37.00 Matches: 8
 Percent Similarity: 88.9% Conservative: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 86.0% Indels: 0
 DB: 2 Gaps: 0
 US-10-774-176-7 (1-9) x AAT75704 (1-2848)
 QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 DB 1013 AGCCCTACTACTCTCTAATGATATCTTA 1039
 RESULT 45
 AAD56099
 ID AAD56099 standard; DNA; 30346 BP.
 AC AAD56099;
 XX
 XX 07-AUG-2003 (first entry)
 DT
 XX Human CCND1 carcinoma associated (CA) gene.
 DE
 XX Carcinoma; gene therapy; carcinoma associated gene; CA; CCND1; human; ds.
 KW
 XX

OS Homo sapiens.
 XX WO2003035837-A2.
 PN
 XX 01-MAY-2003.
 PD
 XX 22-OCT-2002; 2002WO-US033835.
 PF
 XX 23-OCT-2001; 2001US-00004113.
 PR
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 XX Engelhard EK, Morris DW;
 PI WPI; 2003-421412/39.
 XX
 DR New recombinant nucleic acid and its encoded protein, useful for
 XX preparing a composition for diagnosing or treating carcinomas.
 FT
 XX
 PS Claim 1; Page 94-98; 173pp; English.
 XX
 CC The invention relates to novel sequences which are useful for preparing a
 CC composition for diagnosing or treating carcinomas. These sequence are
 CC also useful in gene therapy. The present sequence is human CCND1
 CC carcinoma associated (CA) gene. This sequence is used in the invention
 CC
 XX Sequence 30346 BP; 6484 A; 8431 C; 8729 G; 6702 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1.54e+04 Length: 30346
 Score: 37.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 86.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-7 (1-9) x AAD56099 (1-30346)
 QY 2 LeuGlnThrSerTyrValPheLeu 9
 DB 383 CTTCAACCCAGCTATGTATTATT 406
 RESULT 46
 ADA02461
 ID ADA02461 standard; DNA; 30346 BP.
 XX
 AC ADA02461;
 XX
 DT 06-NOV-2003 (first entry)
 DT
 XX Human CCND1 carcinoma associated gene, SEQ ID NO:980.
 DE
 XX Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
 KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
 KW gene; ds.
 KW
 XX Homo sapiens.
 OS
 XX WO2003057146-A2.
 PN
 XX 17-JUL-2003.
 PD
 XX 26-DEC-2002; 2002WO-US041414.
 PF
 XX 26-DEC-2001; 2001US-00035832.
 PR
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 XX Morris DW;
 PI WPI; 2003-587068/55.
 XX
 DR New recombinant nucleic acid encoding carcinoma associated protein,
 PT

PT useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 980; 245pp; English.

CC The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 30346 BP; 6484 A; 8431 C; 8729 G; 6702 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.54e+04 Length: 30346
Score: 37.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 86.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-7 (1-9) x ADA02461 (1-30346)

QY 2 LeuGlnThrSerTyrValPheLeu 9
|||||
DB 383 CTTCAACACGCTATGTATTATT 406

RESULT 47
ADB72200
ID ADB72200 standard; DNA; 30346 BP.

AC ADB72200;

XX 04-DEC-2003 (first entry)

XX Human CCND1 gene.

XX human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
XX cancer; neoplasm; adenocarcinoma; sarcoma; gene.

XX Homo sapiens.

XX WO2003008583-A2.

XX 30-JAN-2003.

XX 26-DEC-2001; 2001WO-US051291.

XX 02-MAR-2001; 2001US-00798586.

XX 23-OCT-2001; 2001US-00004113.

XX 08-NOV-2001; 2001US-00052482.

XX 30-NOV-2001; 2001US-00997722.

XX 20-DEC-2001; 2001US-00034650.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhard EK;

XX WPI; 2003-239337/23.

XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.

XX Claim 1; SEQ ID NO 28; 2304pp; English.

XX The invention relates to a novel recombinant nucleic acid comprising a nucleotide sequence selected from any of the 660 sequences fully defined in the specification. A polynucleotide of the invention has cytostatic activity, and may have a use in gene therapy, or in a vaccine. The recombinant nucleic acids and polypeptides are useful for treating carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and CC sarcomas. The present sequence represents a human gene of the invention.

XX Sequence 30346 BP; 6484 A; 8431 C; 8729 G; 6702 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.54e+04 Length: 30346
Score: 37.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 86.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ADB72200 (1-30346)

QY 2 LeuGlnThrSerTyrValPheLeu 9
|||||
DB 383 CTTCAACACGCTATGTATTATT 406

RESULT 48
ADB82932
ID ADB82932 standard; DNA; 30346 BP.

XX ADB82932;

XX 29-JAN-2004 (first entry)

XX Human CCND1 genomic DNA sequence.

XX human; cancer-associated nucleic acid; screening; cancer; lymphoma;
XX leukaemia; breast cancer; gene therapy; vaccine; ds.

XX Homo sapiens.

XX WO2003080808-A2.

XX 02-OCT-2003.

XX 21-MAR-2003; 2003WO-US008919.

XX 21-MAR-2002; 2002US-0367025P.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-865119/80.

XX New cancer-associated proteins and nucleic acids, useful for screening
PT for anticancer activity in a potential drug, or for detecting,
PT diagnosing, preventing and treating cancers, e.g. lymphoma, leukemia or
PT breast cancer.

XX Claim 16; SEQ ID NO 16; 248pp; English.

XX The invention comprises human and mouse cancer-associated nucleic acid sequences. The cancer associated nucleic acids of the invention are useful for screening for anticancer activity in a potential drug, as well as detecting, diagnosing, preventing and treating cancers (e.g. lymphoma, CC leukaemia, or breast cancer). The present sequence represents a cancer-

CC associated nucleic acid of the invention.
XX
SQ Sequence 30346 BP; 6484 A; 8431 C; 8729 G; 6702 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.54e+04 Length: 30346
Score: 37.00 Matches: 7
Percent Similarity: 100.0% Conservativity: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 86.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-7 (1-9) x ADE82932 (1-30346)

QY 2 LeuGlnThrSerTyrValPheLeu 9
DB 383 CTTCAACCCAGCTATGTTATT 406

RESULT 49
ABD33387
ID ABD33387 standard; DNA; 176594 BP.

XX
AC ABD33387;
XX
XX 18-NOV-2004 (first entry)
XX
XX Murine cancer-associated (CA) gene MD07-072.

XX
XX Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
XX ds; cancer; cytostatic.

XX Mus musculus.

XX WO2004058146-A2.

XX 15-JUL-2004.

XX 15-DEC-2003; 2003WO-US040081.

XX 17-DEC-2002; 2002US-00322281.

XX (SAGR-) SAGRES DISCOVERY INC.

XX Morris DW, Malandro MS;

XX WPI; 2004-499109/47.

XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.

XX Disclosure; SEQ ID NO 495; 182pp; English.

XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a murine CA gene of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 176594 BP; 46235 A; 37663 C; 39590 G; 52160 T; 0 U; 946 Other;

Alignment Scores:
Pred. No.: 1.12e+05 Length: 176594

Score: 37.00 Matches: 7
Percent Similarity: 88.9% Conservativity: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 86.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-7 (1-9) x ABD33387 (1-176594)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
DB 7637 AGCCTTCAACATCATACATTTTCCTG 7663

RESULT 50
ABL75414
ID ABL75414 standard; cDNA; 199 BP.

XX ABL75414;
XX
XX 14-MAY-2002 (first entry)

XX Corn tassel-derived polynucleotide (cdps) SEQ ID NO:4788.
DE
XX
XX Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassel; gene; ss.

XX Zea mays.

XX US2001051335-A1.

XX 13-DEC-2001.

XX 16-APR-1999; 99US-00294093.

XX 21-APR-1998; 98US-0082567P.

XX (LALG/) LALGUDI R V.
XX (ITOL/) ITO L Y.

XX (SHER/) SHERMAN B K.

XX Lalgudi RV, Ito LY, Sherman BK;

XX WPI; 2002-163647/21.

XX Novel purified corn tassel-derived polynucleotide useful for determining
PT altered gene expression, to recover regulatory elements and to follow
PT inheritance of desirable characteristics through hybrid breeding
PT programs.

XX Claim 1; SEQ ID NO 4788; 201pp; English.

XX The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences
CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful in
CC the evaluation, and alteration of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC quality and yield, and as molecular markers for studying inheritance of
CC multigene traits in a plant breeding program. (I) can be used to produce
CC a tassel-specific profile of gene transcription, a transcript image, to
CC clone regulatory elements for use in transformation vectors, to express a
CC polypeptide, to identify, isolate or extend identical or related corn
CC tassel nucleic acid sequences from DNA libraries, in nucleic acid
CC hybridisation or amplification technologies, as query sequences to
CC determine homology of known sequences, as probe for use in Southern or
CC Northern hybridisation, and to identify the presence of and/or to
CC determine the degree of similarity between two (or more) nucleic acid
CC sequences

SQ Sequence 199 BP; 33 A; 42 C; 51 G; 73 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 87 Length: 199
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 83.7% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x ABL75414 (1-199)

Qy 1 SerLeuGlnThrSerTyrValphe 8

Db 89 AGCCTGCAAACTTCTTACGTTTAT 112

Search completed: April 25, 2006, 12:35:21
Job time : 340.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using framep2n_model
Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignments)
171.290 Million cell updates/sec

Title: US-10-774-176-7

Perfect score: 43

Sequence: 1 SLQTSYVFL 9

Scoring table:

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5893141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Command line parameters:

-MODEL=framep2n_model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10774176/runat_24042006_165114_19197/app_query.fasta_1
-DB=GenEmbl -QFWT=fastap -SUPFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bicsum62 -TRANS=human40 cdi -LIST=1000
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes04
-USER=US10774176 @CCN_1.1 6765 @runat_24042006_165114_19197 -NCFU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_ov.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_hgt.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	ID	Description
1	43	100.0	65	6	CQ559387	Sequence
2	43	100.0	290	6	CQ687716	Sequence
3	43	100.0	475	6	CQ920916	Sequence

4	43	100.0	901	6	BD249733	BD249733 Polypepti
5	43	100.0	901	6	AX025013	AX025013 Sequence
6	43	100.0	901	6	AX316088	AX316088 Sequence
7	43	100.0	927	6	AX829164	AX829164 Sequence
8	43	100.0	1260	6	AX467373	AX467373 Sequence
9	43	100.0	1260	6	AX821533	AX821533 Sequence
10	43	100.0	1260	6	AX821548	AX821548 Sequence
11	43	100.0	1263	6	BD249731	BD249731 Polypepti
12	43	100.0	1263	6	AX025011	AX025011 Sequence
13	43	100.0	1263	6	AX149553	AX149553 Sequence
14	43	100.0	1263	6	AX316086	AX316086 Sequence
15	43	100.0	1263	6	AX467371	AX467371 Sequence
16	43	100.0	1281	6	BD249732	BD249732 Polypepti
17	43	100.0	1281	6	AX025012	AX025012 Sequence
18	43	100.0	1281	6	AX316087	AX316087 Sequence
19	43	100.0	2053	8	HS5740A	Z29083 Homo sapien
20	43	100.0	2183	5	CR855786	CR855786 Xenopus t
21	43	100.0	2333	9	AF063939	AF063939 Rattus no
22	43	100.0	2359	6	BD127282	BD127282 Primer fo
23	43	100.0	2359	6	CQ782724	CQ782724 Sequence
24	43	100.0	2359	8	AK074786	AK074786 Homo sapi
25	43	100.0	2361	6	BD127283	BD127283 Primer fo
26	43	100.0	2361	6	CQ782726	CQ782726 Sequence
27	43	100.0	2361	6	AX961916	AX961916 Sequence
28	43	100.0	2361	8	AK074790	AK074790 Homo sapi
29	43	100.0	2361	9	BC087011	BC087011 Rattus no
30	43	100.0	2379	8	BC037161	BC037161 Homo sapi
31	43	100.0	2423	9	BC058198	BC058198 Mus muscu
32	43	100.0	2557	6	AX961912	AX961912 Sequence
33	43	100.0	2557	6	AX961914	AX961914 Sequence
34	43	100.0	2714	8	AB168308	AB168308 Macaca fa
35	43	100.0	5551	8	HSA012159	HA012159 Homo sapi
36	43	100.0	7942	8	MMU012160	MMU012160 Mus muscu
37	43	100.0	121909	8	HSJ492P14	AL121977 Human DNA
38	43	100.0	167046	9	AC158516	AC158516 Mus muscu
39	43	100.0	210237	14	AC128294	AC128294 Rattus no
40	43	100.0	239076	14	AC106962	AC106962 Rattus no
41	40	93.0	1713	5	BC099984	BC099984 Danio rer
42	40	93.0	1765	5	AB097825	AB097825 Danio rer
43	40	93.0	110000	14	AL353584_2	Continuation (3 of
44	40	93.0	131711	14	AL445194	AL445194 Homo sapi
45	40	93.0	132139	8	AL139114	AL139114 Human DNA
46	40	93.0	178603	9	AC119204	AC119204 Mus muscu
47	40	93.0	188514	9	AC116853	AC116853 Mus muscu
48	40	93.0	213793	8	AC019212	AC019212 Homo sapi
49	39	90.7	64140	14	AC104017	AC104017 Homo sapi
50	39	90.7	110000	15	AP008212_089	Continuation (90 o
51	39	90.7	110373	14	AP004321	AP004321 Oryza sat
52	39	90.7	144495	15	AP003575	AP003575 Oryza sat
53	39	90.7	153335	14	CR854840	CR854840 Danio rer
54	39	90.7	153366	9	AC140347	AC140347 Mus muscu
55	39	90.7	159138	9	AC124403	AC124403 Mus muscu
56	39	90.7	160903	9	AC132260	AC132260 Mus muscu
57	39	90.7	165768	14	AL672011	AL672011 Mus muscu
58	39	90.7	166679	14	AC105366	AC105366 Rattus no
59	39	90.7	171074	5	CR925725	CR925725 Zebrafish
60	39	90.7	172052	9	AL645747	AL645747 Mouse DNA
61	39	90.7	172357	14	AP001808	AP001808 Homo sapi
62	39	90.7	173239	8	AC006478	AC006478 Homo sapi
63	39	90.7	179414	14	AC165320	AC165320 Mus muscu
64	39	90.7	181552	8	AC012417	AC012417 Homo sapi
65	39	90.7	189215	9	AC124490	AC124490 Mus muscu
66	39	90.7	202490	14	AP002426	AP002426 Homo sapi
67	39	90.7	203548	14	AC102094	AC102094 Mus muscu
68	39	90.7	204327	5	BX323822	BX323822 Zebrafish
69	39	90.7	206878	9	AC157667	AC157667 Mus muscu
70	39	90.7	216476	5	BX284112	BX284112 Zebrafish
71	39	90.7	217726	9	AL928680	AL928680 Mouse DNA
72	39	90.7	227595	9	AC111275	AC111275 Rattus no
73	39	90.7	228985	9	AC154911	AC154911 Mus muscu
74	39	90.7	239796	14	AC119389	AC119389 Rattus no
75	39	90.7	242533	14	AC106659	AC106659 Rattus no
76	39	90.7	242867	14	AC097599	AC097599 Rattus no

C 77	39	90.7	254939	14	AC099076	AC099076 Rattus no	C 150	37	86.0	110000	1	BA000019_55	Continuation (56 o
C 78	39	90.7	256453	14	AC120777	AC120777 Rattus no	C 151	37	86.0	110000	14	BX901880_2	Continuation (3 of
C 79	39	90.7	271189	14	AC128515	AC128515 Rattus no	C 152	37	86.0	110000	15	AP008214_051	Continuation (52 o
C 80	39	90.7	349519	1	BX571862	BX571862 Photorhab	C 153	37	86.0	110000	15	AP008214_052	Continuation (53 o
C 81	38	88.4	450	6	AR500353	AR500353 Sequence	C 154	37	86.0	110000	15	AP008217_136	Continuation (137
C 82	38	88.4	450	6	AR515635	AR515635 Sequence	C 155	37	86.0	110000	15	CR382132_01	Continuation (2 of
C 83	38	88.4	743	5	AV550479	AV550479 Salmo sal	C 156	37	86.0	110000	15	AP008209_358	Continuation (359
C 84	38	88.4	4854	14	AC018155	AC018155 Drosophil	C 157	37	86.0	113738	14	AC138298	AC138298 Mus muscu
C 85	38	88.4	86189	5	AC140939	AC140939 Gallus ga	C 158	37	86.0	123430	8	AP001824	AP001824 Homo sapi
C 86	38	88.4	95345	8	AC007029	AC007029 Homo sapi	C 159	37	86.0	125565	9	AC1359762	AC1359762 Human DNA
C 87	38	88.4	106601	8	AL357352	AL357352 Human DNA	C 160	37	86.0	139563	9	AC132837	AC132837 Mus muscu
C 88	38	88.4	110000	14	TAMN4_10	Continuation (11 o	C 161	37	86.0	139759	14	AC147778	AC147778 Mus muscu
C 89	38	88.4	117016	14	AC151655	AC151655 Dasyuon	C 162	37	86.0	140920	15	AC146947	AC146947 Oryza sat
C 90	38	88.4	134687	8	AL445259	AL445259 Human DNA	C 163	37	86.0	141620	14	AC068008	AC068008 Homo sapi
C 91	38	88.4	150235	8	AC147100	AC147100 Pan trogl	C 164	37	86.0	143072	8	HSJ415P21	AL117345 Human DNA
C 92	38	88.4	150313	8	AC137672	AC137672 Homo sapi	C 165	37	86.0	143796	14	AC006411	AC006411 Homo sapi
C 93	38	88.4	152064	5	BX510349	BX510349 Zebrafish	C 166	37	86.0	144211	5	BX322618	BX322618 Zebrafish
C 94	38	88.4	156107	8	AC099543	AC099543 Homo sapi	C 167	37	86.0	148549	14	AL354818	AL354818 Homo sapi
C 95	38	88.4	176328	8	AC146114	AC146114 Pan trogl	C 168	37	86.0	149527	8	AC092504	AC092504 Homo sapi
C 96	38	88.4	176493	8	AC102661	AC102661 Mus muscu	C 169	37	86.0	150355	8	HSJ364H10	AL078603 Human DNA
C 97	38	88.4	179148	8	AC145967	AC145967 Pan trogl	C 170	37	86.0	150900	8	AC068075	AC068075 Homo sapi
C 98	38	88.4	179691	8	AC006572	AC006572 Homo sapi	C 171	37	86.0	152649	14	AC018353	AC018353 Homo sapi
C 99	38	88.4	180640	14	AC147946	AC147946 Macropus	C 172	37	86.0	155137	14	CR769771	CR769771 Danto rer
C 100	38	88.4	181143	14	AC120652	AC120652 Rattus no	C 173	37	86.0	158560	8	AC024614	AC024614 Homo sapi
C 101	38	88.4	183457	14	AC128106	AC128106 Rattus no	C 174	37	86.0	160901	14	AC120060	AC120060 Rattus no
C 102	38	88.4	182265	8	AC146204	AC146204 Pan trogl	C 175	37	86.0	162560	5	AL954384	AL954384 Zebrafish
C 103	38	88.4	194022	14	AC164937	AC164937 Oryctolag	C 176	37	86.0	162787	9	AC140350	AC140350 Mus muscu
C 104	38	88.4	194520	8	AL356915	AL356915 Human DNA	C 177	37	86.0	169313	8	AP001782	AP001782 Homo sapi
C 105	38	88.4	199087	9	AC102340	AC102340 Mus muscu	C 178	37	86.0	169690	9	AL929373	AL929373 Mouse DNA
C 106	38	88.4	201990	8	AC002385	AC002385 Homo sapi	C 179	37	86.0	170849	5	CR405715	CR405715 Zebrafish
C 107	38	88.4	205242	8	AC146375	AC146375 Pan trogl	C 180	37	86.0	173656	8	AC135113	AC135113 Mus muscu
C 108	38	88.4	205650	14	AC160470	AC160470 Mus muscu	C 181	37	86.0	173701	8	AC018946	AC018946 Homo sapi
C 109	38	88.4	210536	14	AC165158	AC165158 Mus muscu	C 182	37	86.0	176493	9	AC102861	AC102861 Mus muscu
C 110	38	88.4	213660	14	AC073699	AC073699 Mus muscu	C 183	37	86.0	176969	8	AC127702	AC127702 Homo sapi
C 111	38	88.4	219865	5	BX469919	BX469919 Zebrafish	C 184	37	86.0	178211	14	AC120657	AC120657 Rattus no
C 112	38	88.4	220238	14	AC163639	AC163639 Mus muscu	C 185	37	86.0	179744	8	AC008133	AC008133 Homo sapi
C 113	38	88.4	224010	14	AC117300	AC117300 Rattus no	C 186	37	86.0	179844	8	AC009265	AC009265 Homo sapi
C 114	38	88.4	228373	14	AC095267	AC095267 Rattus no	C 187	37	86.0	181573	14	AC020959	AC020959 Mus muscu
C 115	38	88.4	229491	14	AC151219	AC151219 Bos tauru	C 188	37	86.0	181900	9	BX284114	BX284114 Mouse DNA
C 116	38	88.4	233425	9	AC099722	AC099722 Mus muscu	C 189	37	86.0	182452	8	AL161731	AL161731 Human DNA
C 117	38	88.4	234152	14	AC111434	AC111434 Rattus no	C 190	37	86.0	183252	14	AC023637	AC023637 Homo sapi
C 118	38	88.4	238157	14	AC164703	AC164703 Mus muscu	C 191	37	86.0	185916	14	AP001568	AP001568 Homo sapi
C 119	38	88.4	247900	14	AC107565	AC107565 Rattus no	C 192	37	86.0	186036	15	AP005531	AP005531 Oryza sat
C 120	38	88.4	248154	14	AC111496	AC111496 Rattus no	C 193	37	86.0	187010	9	AC102472	AC102472 Mus muscu
C 121	38	88.4	265501	14	AC112367	AC112367 Rattus no	C 194	37	86.0	188779	8	AC090701	AC090701 Homo sapi
C 122	38	88.4	266759	14	AC156669	AC156669 Bos tauru	C 195	37	86.0	190749	14	AF321233	AF321233 Mus muscu
C 123	38	88.4	284329	14	AC163423	AC163423 Bos tauru	C 196	37	86.0	190822	14	AC051613	AC051613 Mus muscu
C 124	38	88.4	311391	14	AC151079	AC151079 Bos tauru	C 197	37	86.0	195025	9	MMCMX137	AL021127 Mus muscu
C 125	38	88.4	349082	1	BX572091	BX572091 Prochlozo	C 198	37	86.0	195446	9	AC133967	AC133967 Mus muscu
C 126	37	86.0	591	10	G76784	G76784 S208F6667FH	C 199	37	86.0	196155	5	BX247882	BX247882 Zebrafish
C 127	37	86.0	606	10	BV061074	BV061074 S212P6009	C 200	37	86.0	197165	9	AC151907	AC151907 Mus muscu
C 128	37	86.0	711	10	BV015274	BV015274 S212P6037	C 201	37	86.0	199768	14	AP001589	AP001589 Homo sapi
C 129	37	86.0	718	10	BV620259	BV620259 S217P6848	C 202	37	86.0	200707	1	BSUB0015	Z99118 Bacillus su
C 130	37	86.0	1251	9	BC099186	BC099186 Rattus no	C 203	37	86.0	201076	14	AC119349	AC119349 Rattus su
C 131	37	86.0	5010	8	HSR50809	HSR50809 Homo sapi	C 204	37	86.0	201511	14	AC154641	AC154641 Mus muscu
C 132	37	86.0	6850	15	AR193555	AR193555 Lactuca s	C 205	37	86.0	203405	14	AC024964	AC024964 Homo sapi
C 133	37	86.0	30346	6	AX695353	AX695353 Sequence	C 206	37	86.0	205224	9	AL671908	AL671908 Mouse DNA
C 134	37	86.0	67593	14	AC079197	AC079197 Homo sapi	C 207	37	86.0	208153	9	AL627188	AL627188 Mouse DNA
C 135	37	86.0	72519	14	AC087458	AC087458 Homo sapi	C 208	37	86.0	212060	14	AC148416	AC148416 Callithri
C 136	37	86.0	76880	14	AC150175	AC150175 Gallus ga	C 209	37	86.0	214591	14	AC114131	AC114131 Rattus no
C 137	37	86.0	77242	5	AL591462	AL591462 Zebrafish	C 210	37	86.0	216837	14	AC126597	AC126597 Rattus no
C 138	37	86.0	83715	14	AC138269	AC138269 Mus muscu	C 211	37	86.0	218811	14	AC159287	AC159287 Mus muscu
C 139	37	86.0	85767	5	BX537309	BX537309 Zebrafish	C 212	37	86.0	219538	14	AC162272	AC162272 Bos tauru
C 140	37	86.0	92392	15	AC128647	AC128647 Oryza sat	C 213	37	86.0	220060	1	AF008220	AF008220 Bacillus
C 141	37	86.0	93432	8	QNS01DSR	AL121840 Human chr	C 214	37	86.0	221393	14	AC159508	AC159508 Bos tauru
C 142	37	86.0	93832	14	AP008176	AP008176 Lotus cor	C 215	37	86.0	221523	14	AC157389	AC157389 Bos tauru
C 143	37	86.0	96428	2	AC104269	AC104269 Caenorhab	C 216	37	86.0	221545	9	AL844855	AL844855 Mouse DNA
C 144	37	86.0	99128	8	AC007948	AC007948 Genomic s	C 217	37	86.0	221849	5	BX293568	BX293568 Zebrafish
C 145	37	86.0	102137	5	AL591374	AL591374 Zebrafish	C 218	37	86.0	223972	14	AC027478	AC027478 Homo sapi
C 146	37	86.0	104138	15	AC151804	AC151804 Solanum d	C 219	37	86.0	225053	14	AC084396	AC084396 Homo sapi
C 147	37	86.0	106179	8	HS790B6	AL031677 Human DNA	C 220	37	86.0	225122	9	AC099710	AC099710 Mus muscu
C 148	37	86.0	107641	9	AF125313	AF125313 Mus muscu	C 221	37	86.0	226657	9	AL731843	AL731843 Mouse DNA
C 149	37	86.0	109642	8	AC119734	AC119734 Homo sapi	C 222	37	86.0	228044	5	BX682552	BX682552 Zebrafish

C 223	37	86.0	232058	14	AC118324	Rattus no	AC118324	Rattus no	296	36	83.7	103202	14	AC087107	AC087107 Homo sapi
C 224	37	86.0	235419	14	AC095979	Rattus no	AC095979	Rattus no	297	36	83.7	104878	15	AC121238	AC121238 Medicago
C 225	37	86.0	240932	14	AC098206	Rattus no	AC098206	Rattus no	298	36	83.7	110000	1	BA000011_12	Continuation (13 o
C 226	37	86.0	242126	9	AC115629	Mus muscu	AC115629	Mus muscu	299	36	83.7	110000	1	BA000011_13	Continuation (14 o
C 227	37	86.0	250840	14	AC153148	Mus muscu	AC153148	Mus muscu	C 300	36	83.7	110000	14	AC125752_0	AC125752 Rattus no
C 228	37	86.0	253261	14	AC109368	Rattus no	AC109368	Rattus no	C 301	36	83.7	110000	14	AC129397_0	AC129397 Rattus no
C 229	37	86.0	253824	9	AC137524	Mus muscu	AC137524	Mus muscu	C 302	36	83.7	110000	15	AP008214_090	Continuation (91 o
C 230	37	86.0	255430	14	AC099387	Rattus no	AC099387	Rattus no	C 303	36	83.7	110000	15	AP008217_188	Continuation (189
C 231	37	86.0	256720	9	AC110241	Mus muscu	AC110241	Mus muscu	C 304	36	83.7	110000	15	AP008217_189	Continuation (190
C 232	37	86.0	264321	9	AL080762	Mouse DNA	AL080762	Mouse DNA	C 305	36	83.7	110000	15	AP008219_12	Continuation (13 o
C 233	37	86.0	268123	9	AC093479	Mus muscu	AC093479	Mus muscu	C 306	36	83.7	110000	15	AP008207_235	Continuation (236
C 234	37	86.0	280559	14	AC157156	Bos tauru	AC157156	Bos tauru	C 307	36	83.7	110000	15	AP008207_236	Continuation (237
C 235	37	86.0	317508	14	AC106387	Rattus no	AC106387	Rattus no	C 308	36	83.7	110565	15	AC009243	AC009243 Genomic s
C 236	36	83.7	148	1	RP2251262				C 309	36	83.7	111074	8	AC007569	AC007569 Homo sapi
C 237	36	83.7	270	6	AX478373	Sequence	AX478373	Sequence	C 310	36	83.7	112182	14	AC141177	AC141177 Rattus no
C 238	36	83.7	502	1	AF184910	Bradyrhiz	AF184910	Bradyrhiz	C 311	36	83.7	112390	8	AC016607	AC016607 Homo sapi
C 239	36	83.7	519	1	AF239246	Bradyrhiz	AF239246	Bradyrhiz	C 312	36	83.7	112765	14	AC012129	AC012129 Homo sapi
C 240	36	83.7	607	10	BV001986	S208P6521	BV001986	S208P6521	C 313	36	83.7	113085	8	AC006150	AC006150 Homo sapi
C 241	36	83.7	622	10	BV388377	S244P618R	BV388377	S244P618R	C 314	36	83.7	115194	5	BX569782	BX569782 Zebrafish
C 242	36	83.7	647	10	BV243796	S234P6168	BV243796	S234P6168	C 315	36	83.7	116265	5	AC140941	AC140941 Gallus ga
C 243	36	83.7	675	10	BV456689	qrl5Shil.	BV456689	qrl5Shil.	C 316	36	83.7	117559	8	AC092279	AC092279 Homo sapi
C 244	36	83.7	702	10	BV456880	S216P6211	BV456880	S216P6211	C 317	36	83.7	119814	8	AL606535	AL606535 Human DNA
C 245	36	83.7	824	5	CR760872	Xenopus t	CR760872	Xenopus t	C 318	36	83.7	120376	14	AC145027	AC145027 Medicago
C 246	36	83.7	824	6	BD100097	Novel gen	BD100097	Novel gen	C 319	36	83.7	123413	8	AL138834	AL138834 Human DNA
C 247	36	83.7	824	6	BD020159	Novel gen	BD020159	Novel gen	C 320	36	83.7	123532	8	AC092881	AC092881 Homo sapi
C 248	36	83.7	836	6	AR088267	Sequence	AR088267	Sequence	C 321	36	83.7	124156	8	AL596132	AL596132 Human DNA
C 249	36	83.7	836	6	AR097298	Sequence	AR097298	Sequence	C 322	36	83.7	124660	14	AP001954	AP001954 Homo sapi
C 250	36	83.7	893	10	BV461036	qcd56c12.	BV461036	qcd56c12.	C 323	36	83.7	125913	14	AC147526	AC147526 Oroleumur
C 251	36	83.7	899	5	BC063337	Xenopus t	BC063337	Xenopus t	C 324	36	83.7	126382	14	AC008879	AC008879 Homo sapi
C 252	36	83.7	1155	6	AX478374	Sequence	AX478374	Sequence	C 325	36	83.7	128119	15	AC140721	AC140721 Medicago
C 253	36	83.7	1213	1	NW4J5010				C 326	36	83.7	131181	14	CT010459	CT010459 Medicago
C 254	36	83.7	1215	1	NSDAJ5008				C 327	36	83.7	131747	14	CT009479	CT009479 Medicago
C 255	36	83.7	1262	1	AF338178	Rhodopseu	AF338178	Rhodopseu	C 328	36	83.7	134151	8	AC079318	AC079318 Homo sapi
C 256	36	83.7	1294	15	AK111788	Oryza sat	AK111788	Oryza sat	C 329	36	83.7	134764	8	AP003305	AP003305 Homo sapi
C 257	36	83.7	1297	1	RP2J5013	Rhodopseu	RP2J5013	Rhodopseu	C 330	36	83.7	140682	8	HS283K11	HS283K11 Human DNA
C 258	36	83.7	1400	1	AF338168	Bradyrhiz	AF338168	Bradyrhiz	C 331	36	83.7	140697	14	AC021289	AC021289 Homo sapi
C 259	36	83.7	1485	1	AF338172	Bradyrhiz	AF338172	Bradyrhiz	C 332	36	83.7	141363	14	AY245861	AY245861 Homo sapi
C 260	36	83.7	1538	1	AF338169	Bradyrhiz	AF338169	Bradyrhiz	C 333	36	83.7	142094	8	HS0543J19	HS0543J19 Human DNA
C 261	36	83.7	1564	2	AK115992	Ciona int	AK115992	Ciona int	C 334	36	83.7	142420	8	AL138816	AL138816 Human DNA
C 262	36	83.7	1578	1	AF338173	Bradyrhiz	AF338173	Bradyrhiz	C 335	36	83.7	142978	8	AC097717	AC097717 Homo sapi
C 263	36	83.7	1590	1	AF338160	Bradyrhiz	AF338160	Bradyrhiz	C 336	36	83.7	143777	14	AC147817	AC147817 Xenopus t
C 264	36	83.7	2101	10	BV177018	sqm92202	BV177018	sqm92202	C 337	36	83.7	145168	8	AC105231	AC105231 Homo sapi
C 265	36	83.7	2179	15	BT018733	Zea mays	BT018733	Zea mays	C 338	36	83.7	145631	14	AC156356	AC156356 Carolia
C 266	36	83.7	2444	6	AR578462	Sequence	AR578462	Sequence	C 339	36	83.7	146688	8	AC022261	AC022261 Homo sapi
C 267	36	83.7	2526	8	BC010649	Homo sapi	BC010649	Homo sapi	C 340	36	83.7	146937	14	AC112817	AC112817 Rattus no
C 268	36	83.7	2528	6	C0723812	Sequence	C0723812	Sequence	C 341	36	83.7	147875	14	AC013520	AC013520 Homo sapi
C 269	36	83.7	2535	8	AF151865	Homo sapi	AF151865	Homo sapi	C 342	36	83.7	148503	8	AL135908	AL135908 Human DNA
C 270	36	83.7	2589	8	BC013969	Homo sapi	BC013969	Homo sapi	C 343	36	83.7	148829	15	AC134925	AC134925 Oryza sat
C 271	36	83.7	2826	1	RP23SRRN				C 344	36	83.7	150840	9	AY151050	AY151050 Mus muscu
C 272	36	83.7	2840	1	DQ0111942	Bradyrhiz	DQ0111942	Bradyrhiz	C 345	36	83.7	150989	8	AC091922	AC091922 Homo sapi
C 273	36	83.7	2856	1	DQ011939	Bradyrhiz	DQ011939	Bradyrhiz	C 346	36	83.7	151130	9	AC127432	AC127432 Mus muscu
C 274	36	83.7	3000	1	AF338176	Blastobac	AF338176	Blastobac	C 347	36	83.7	151144	5	BX323575	BX323575 Zebrafish
C 275	36	83.7	5076	1	AF338177	Alfipa fe.	AF338177	Alfipa fe.	C 348	36	83.7	152086	8	AL355796	AL355796 Human DNA
C 276	36	83.7	6297	1	AF184625	Rhodopseu	AF184625	Rhodopseu	C 349	36	83.7	152885	8	AC008818	AC008818 Homo sapi
C 277	36	83.7	6364	1	AF184626	Rhodopseu	AF184626	Rhodopseu	C 350	36	83.7	152841	14	AL929463	AL929463 Danio rer
C 278	36	83.7	12932	6	AX645933	Sequence	AX645933	Sequence	C 351	36	83.7	154252	8	AC004062	AC004062 Homo sapi
C 279	36	83.7	12932	8	AB065490	Homo sapi	AB065490	Homo sapi	C 352	36	83.7	154455	15	AP0021844	AP0021844 Homo sapi
C 280	36	83.7	14284	13	AP439352	Choriston	AP439352	Choriston	C 353	36	83.7	154541	15	AP003278	AP003278 Oryza sat
C 281	36	83.7	27893	15	SPAC1039				C 354	36	83.7	155394	14	AC092328	AC092328 Homo sapi
C 282	36	83.7	28871	2	CEC47B2				C 355	36	83.7	155435	14	CR974437	CR974437 Danio rer
C 283	36	83.7	29743	8	AF224492	Homo sapi	AF224492	Homo sapi	C 356	36	83.7	155495	8	AC008825	AC008825 Homo sapi
C 284	36	83.7	31624	2	CBP0208				C 357	36	83.7	156102	14	AC16081	AC16081 Homo sapi
C 285	36	83.7	38569	2	AF016442	Caenorhab	AF016442	Caenorhab	C 358	36	83.7	156394	8	AC105756	AC105756 Homo sapi
C 286	36	83.7	38597	2	AF016442	Caenorhab	AF016442	Caenorhab	C 359	36	83.7	156495	14	AL157374	AL157374 Homo sapi
C 287	36	83.7	42101	8	AC130405				C 360	36	83.7	156899	14	AC103884	AC103884 Homo sapi
C 288	36	83.7	43218	14	AC012979				C 361	36	83.7	157905	15	AP005495	AP005495 Oryza sat
C 289	36	83.7	57261	14	AC090245				C 362	36	83.7	158119	9	AC079098	AC079098 Homo sapi
C 290	36	83.7	63910	14	AC155919	Medicago	AC155919	Medicago	C 363	36	83.7	158411	9	AC154405	AC154405 Mus muscu
C 291	36	83.7	68914	14	AC084144	Homo sapi	AC084144	Homo sapi	C 364	36	83.7	158563	9	AL113446	AL113446 Mus muscu
C 292	36	83.7	75698	14	CR931980_3	Continuation (4 of			C 365	36	83.7	159117	8	AL445426	AL445426 Human DNA
C 293	36	83.7	78608	14	AC022428	Homo sapi	AC022428	Homo sapi	C 366	36	83.7	159215	8	AL360179	AL360179 Human DNA
C 294	36	83.7	94891	5	AC022428	Homo sapi	AC022428	Homo sapi	C 367	36	83.7	159627	5	BX255895	BX255895 Zebrafish
C 295	36	83.7	103171	5	CR3922335				C 368	36	83.7	161462	14	AC093125	AC093125 Papio anu

C 369	83.7	162099	14	AC015730	AC015730 Homo sapi	C 442	83.7	192640	9	AC135354	AC135354 Mus muscu
C 370	83.7	162933	14	AC130433	AC130433 Homo sapi	443	83.7	193390	14	AC112168	AC112168 Homo sapi
C 371	83.7	163163	8	AC069528	AC069528 Homo sapi	444	83.7	194229	14	CR391980	CR391980 Dario rer
C 372	83.7	163277	14	AC084687	AC084687 Homo sapi	445	83.7	194270	8	AC106768	AC106768 Homo sapi
C 373	83.7	163468	14	AC015999	AC015999 Homo sapi	C 446	83.7	194339	5	CR786582	CR786582 Zebrafish
C 374	83.7	163530	14	AC134958	AC134958 Rattus no	C 447	83.7	195661	14	AC116891	AC116891 Mus muscu
C 375	83.7	163593	14	AC018876	AC018876 Homo sapi	C 448	83.7	195691	14	AC017062	AC017062 Homo sapi
C 376	83.7	164293	8	AC091135	AC091135 Homo sapi	C 449	83.7	197156	15	AC133248	AC133248 Oryza sat
C 377	83.7	166207	8	AC009757	AC009757 Homo sapi	C 450	83.7	197577	14	AC121085	AC121085 Mus muscu
C 378	83.7	166918	8	AC009757	AC009757 Homo sapi	C 451	83.7	198349	9	AC123557	AC123557 Mus muscu
C 379	83.7	167248	9	AC102297	AC102297 Mus muscu	C 452	83.7	198671	14	AC136001	AC136001 Homo sapi
C 380	83.7	167283	14	AC011612	AC011612 Homo sapi	C 453	83.7	198733	14	AC166310	AC166310 Oryctolag
C 381	83.7	167441	14	AC092528	AC092528 Papio anu	C 454	83.7	199910	5	AL928725	AL928725 Zebrafish
C 382	83.7	167470	14	AC136060	AC136060 Rattus no	C 455	83.7	200131	9	AC146977	AC146977 Mus muscu
C 383	83.7	168495	9	AC157098	AC157098 Mus muscu	C 456	83.7	200673	9	AC131676	AC131676 Mus muscu
C 384	83.7	168698	8	AC068138	AC068138 Homo sapi	C 457	83.7	200922	9	AC123856	AC123856 Mus muscu
C 385	83.7	169417	14	CR753839	CR753839 Dario rer	C 458	83.7	200965	14	AC006740	AC006740 Caenorhab
C 386	83.7	169542	14	AL356482	AL356482 Homo sapi	C 459	83.7	202405	8	AC109486	AC109486 Homo sapi
C 387	83.7	169611	14	AC010845	AC010845 Drosophil	C 460	83.7	203000	9	AL596130	AL596130 Mouse DNA
C 388	83.7	171362	14	AC150167	AC150167 Gallus ga	C 461	83.7	204605	14	AL160259	AL160259 Homo sapi
C 389	83.7	171534	14	AC152458	AC152458 Carollia	C 462	83.7	204748	14	AC165885	AC165885 Bos tauru
C 390	83.7	172364	9	AL954189	AL954189 Mouse DNA	C 463	83.7	205393	14	EX088574	EX088574 Dario rer
C 391	83.7	172512	8	AC016903	AC016903 Homo sapi	C 464	83.7	205527	14	AC117875	AC117875 Rattus no
C 392	83.7	172786	14	AC142378	AC142378 Rattus no	C 465	83.7	206196	14	AC135295	AC135295 Rattus no
C 393	83.7	173062	14	AC140637	AC140637 Macaca mu	C 466	83.7	206288	9	AC131758	AC131758 Mus muscu
C 394	83.7	173343	9	AC116861	AC116861 Mus muscu	C 467	83.7	206511	14	AC137798	AC137798 Homo sapi
C 395	83.7	173515	9	AC156632	AC156632 Mus muscu	C 468	83.7	206842	9	AL844208	AL844208 Mouse DNA
C 396	83.7	174224	9	AC123868	AC123868 Mus muscu	C 469	83.7	207154	14	AC114180	AC114180 Rattus no
C 397	83.7	174550	14	AC119378	AC119378 Rattus no	C 470	83.7	207748	14	AC154655	AC154655 Mus muscu
C 398	83.7	175302	14	AC079350	AC079350 Homo sapi	C 471	83.7	208335	14	AC164130	AC164130 Bos tauru
C 399	83.7	176273	8	AL135790	AL135790 Human DNA	C 472	83.7	209710	14	AC150051	AC150051 Gallus ga
C 400	83.7	176292	14	AC118151	AC118151 Rattus no	C 473	83.7	211635	14	AC112763	AC112763 Rattus no
C 401	83.7	176547	8	AC092447	AC092447 Homo sapi	C 474	83.7	211852	14	AC152253	AC152253 Bos tauru
C 402	83.7	176797	8	AC116552	AC116552 Homo sapi	C 475	83.7	212126	14	AC128610	AC128610 Rattus no
C 403	83.7	177014	8	AC004063	AC004063 Homo sapi	C 476	83.7	212733	14	AC162019	AC162019 Bos tauru
C 404	83.7	177969	8	AC068107	AC068107 Homo sapi	C 477	83.7	212895	14	AC148788	AC148788 Otollemur
C 405	83.7	178743	14	AC152546	AC152546 Rattus no	C 478	83.7	213991	14	AC130620	AC130620 Rattus no
C 406	83.7	179526	14	AC034133	AC034133 Homo sapi	C 479	83.7	214856	9	AC134434	AC134434 Mus muscu
C 407	83.7	180073	14	AC073096	AC073096 Homo sapi	C 480	83.7	215010	14	AC128614	AC128614 Rattus no
C 408	83.7	180105	8	AC015756	AC015756 Homo sapi	C 481	83.7	216194	9	AC096051	AC096051 Rattus no
C 409	83.7	180331	8	AC007375	AC007375 Homo sapi	C 482	83.7	216379	14	AC103263	AC103263 Rattus no
C 410	83.7	180607	14	AC149243	AC149243 Otollemur	C 483	83.7	216887	14	AC113676	AC113676 Rattus no
C 411	83.7	180727	14	AC027221	AC027221 Homo sapi	C 484	83.7	218147	14	AC137352	AC137352 Rattus no
C 412	83.7	180952	14	AC163228	AC163228 Bos tauru	C 485	83.7	219041	14	AC128323	AC128323 Rattus no
C 413	83.7	181119	14	AC092250	AC092250 Canis fam	C 486	83.7	222048	14	AC095529	AC095529 Rattus no
C 414	83.7	182312	14	AC087709	AC087709 Homo sapi	C 487	83.7	222468	14	AL079525	AL079525 Mus muscu
C 415	83.7	182564	14	AC150115	AC150115 Gallus ga	C 488	83.7	222615	5	AL929345	AL929345 Zebrafish
C 416	83.7	183075	5	AC145959	AC145959 Gallus ga	C 489	83.7	222899	14	AC112416	AC112416 Rattus no
C 417	83.7	183316	2	AC008255	AC008255 Drosophil	C 490	83.7	223110	9	AC123752	AC123752 Mus muscu
C 418	83.7	183436	8	AC012499	AC012499 Homo sapi	C 491	83.7	223935	14	AC135900	AC135900 Rattus no
C 419	83.7	183617	14	AC007524	AC007524 Homo sapi	C 492	83.7	224009	14	AC131058	AC131058 Mus muscu
C 420	83.7	183762	14	AC128232	AC128232 Rattus no	C 493	83.7	224529	14	AC112045	AC112045 Rattus no
C 421	83.7	184635	8	AC025445	AC025445 Homo sapi	C 494	83.7	225492	14	AC160104	AC160104 Mus muscu
C 422	83.7	185736	14	AC135601	AC135601 Homo sapi	C 495	83.7	225541	14	AC126700	AC126700 Rattus no
C 423	83.7	185810	14	AC136052	AC136052 Rattus no	C 496	83.7	225545	14	AC110659	AC110659 Rattus no
C 424	83.7	186451	9	AC122211	AC122211 Mus muscu	C 497	83.7	226374	14	AC110150	AC110150 Rattus no
C 425	83.7	186495	14	AC147081	AC147081 Pan trogl	C 498	83.7	226580	14	AC108446	AC108446 Rattus no
C 426	83.7	186527	9	AC147228	AC147228 Mus muscu	C 499	83.7	226674	14	AC131171	AC131171 Rattus no
C 427	83.7	187279	8	AC113166	AC113166 Homo sapi	C 500	83.7	227214	1	BX572608	BX572608 Rhodopeu
C 428	83.7	187449	8	AC016450	AC016450 Homo sapi	C 501	83.7	227556	14	AC155095	AC155095 Bos tauru
C 429	83.7	187728	14	AC062011	AC062011 Homo sapi	C 502	83.7	228495	14	AC119308	AC119308 Rattus no
C 430	83.7	188389	14	AC024199	AC024199 Homo sapi	C 503	83.7	229599	9	AC140339	AC140339 Mus muscu
C 431	83.7	188464	9	AL845440	AL845440 Mouse DNA	C 504	83.7	230827	8	AC010102	AC010102 Homo sapi
C 432	83.7	188634	9	AC121131	AC121131 Mus muscu	C 505	83.7	231412	14	AC129647	AC129647 Rattus no
C 433	83.7	189263	14	AC137278	AC137278 Rattus no	C 506	83.7	231944	14	AC119630	AC119630 Rattus no
C 434	83.7	189456	8	AC079905	AC079905 Homo sapi	C 507	83.7	232077	14	AC091348	AC091348 Rattus no
C 435	83.7	190927	14	AC141748	AC141748 Apis mell	C 508	83.7	233944	14	AC106583	AC106583 Rattus no
C 436	83.7	191652	8	AL354668	AL354668 Human DNA	C 509	83.7	234223	14	AC099143	AC099143 Rattus no
C 437	83.7	191966	8	AC150280	AC150280 Pan trogl	C 510	83.7	234420	14	AC130023	AC130023 Rattus no
C 438	83.7	192001	14	AC019139	AC019139 Homo sapi	C 511	83.7	234834	9	AC124775	AC124775 Mus muscu
C 439	83.7	192029	5	BX537134	BX537134 Zebrafish	C 512	83.7	234963	14	AC094843	AC094843 Rattus no
C 440	83.7	192353	14	AC021190	AC021190 Homo sapi	C 513	83.7	236078	14	AC106369	AC106369 Rattus no
C 441	83.7	192433	14	AC161272	AC161272 Mus muscu	C 514	83.7	236557	14	AC133972	AC133972 Rattus no
	83.7	192593	9	AC144927	AC144927 Mus muscu						

C 515	36	83.7	237178	14	AC136754	AC136754 Mus muscu	588	35	81.4	684	15	BT014828	BT014828 Arabidops
C 516	36	83.7	237181	14	AC023460	AC023460 Homo sapi	589	35	81.4	705	10	BV603308	BV603308 S217P6706
C 517	36	83.7	237293	14	AC105553	AC105553 Rattus no	590	35	81.4	774	10	BV578547	BV578547 GS91P6146
C 518	36	83.7	238252	14	AC095633	AC095633 Rattus no	C 591	35	81.4	775	10	BV561621	BV561621 Qnc86607.
C 519	36	83.7	238312	14	AC099456	AC099456 Rattus no	C 592	35	81.4	775	10	BV041832	BV041832 S212P6401
C 520	36	83.7	238650	14	AC135943	AC135943 Rattus no	C 593	35	81.4	789	10	BV324541	BV324541 S241P692F
C 521	36	83.7	239438	14	AC107536	AC107536 Rattus no	C 594	35	81.4	794	10	BV589977	BV589977 GS91P6357
C 522	36	83.7	239732	14	AC115490	AC115490 Rattus no	C 595	35	81.4	812	15	BT012597	BT012597 Arabidops
C 523	36	83.7	239772	14	AC125647	AC125647 Rattus no	C 596	35	81.4	838	10	BV566382	BV566382 Qcl98804.
C 524	36	83.7	240092	14	AC159290	AC159290 Mus muscu	C 597	35	81.4	841	15	AY070741	AY070741 Arabidops
C 525	36	83.7	240971	14	AC120956	AC120956 Rattus no	C 598	35	81.4	901	10	BV566033	BV566033 Qc111907.
C 526	36	83.7	241300	14	AC098306	AC098306 Rattus no	C 599	35	81.4	918	10	BV572132	BV572132 GS91P6331
C 527	36	83.7	241643	14	AC162063	AC162063 Bos tauru	600	35	81.4	973	10	BV576882	BV576882 GS91P6243
C 528	36	83.7	242527	14	AC115266	AC115266 Rattus no	601	35	81.4	1138	5	CR339034	CR339034 Gallus ga
C 529	36	83.7	242527	14	AC150646	AC150646 Bos tauru	602	35	81.4	1293	6	AR378552	AR378552 Sequence
C 530	36	83.7	243348	14	AC150646	AC150646 Bos tauru	603	35	81.4	1379	2	AF332521	AF332521 Eimeria a
C 531	36	83.7	243440	14	AC098290	AC098290 Rattus no	C 604	35	81.4	1898	15	BT001215	BT001215 Arabidops
C 532	36	83.7	244722	14	AC111488	AC111488 Rattus no	C 605	35	81.4	2000	6	AX655861	AX655861 Sequence
C 533	36	83.7	245636	14	AC095871	AC095871 Rattus no	C 606	35	81.4	2023	15	BT013731	BT013731 Lycopersi
C 534	36	83.7	246602	14	AC103062	AC103062 Rattus no	C 607	35	81.4	2059	2	AF387489	AF387489 Aedes aeg
C 535	36	83.7	247169	14	AC123007	AC123007 Rattus no	C 608	35	81.4	2087	15	AY081314	AY081314 Arabidops
C 536	36	83.7	247878	14	AC094450	AC094450 Rattus no	C 609	35	81.4	2164	1	AB088041	AB088041 Vibrio pa
C 537	36	83.7	247976	14	AC162598	AC162598 Bos tauru	C 610	35	81.4	2171	15	AY056174	AY056174 Arabidops
C 538	36	83.7	249287	14	AC161834	AC161834 Bos tauru	C 611	35	81.4	2517	2	AB210327	AB210327 Ciona int
C 539	36	83.7	250684	14	AC096960	AC096960 Rattus no	C 612	35	81.4	2518	2	AK174811	AK174811 Ciona int
C 540	36	83.7	251441	14	AC130747	AC130747 Rattus no	C 613	35	81.4	2580	13	TYDMONOCOT	N811023 Tobacco yel
C 541	36	83.7	251570	14	AC094463	AC094463 Rattus no	C 614	35	81.4	2746	15	AA1LV	AY222772 Arxula ad
C 542	36	83.7	252115	14	AC131854	AC131854 Rattus no	C 615	35	81.4	2839	9	AY083459	AY083459 Rattus no
C 543	36	83.7	252185	14	AC122980	AC122980 Rattus no	C 616	35	81.4	2840	15	CAU15152	U15152 Candida alb
C 544	36	83.7	252330	14	AC096167	AC096167 Rattus no	C 617	35	81.4	3475	9	BC034743	BC034743 Mus muscu
C 545	36	83.7	253201	14	AC106168	AC106168 Rattus no	C 618	35	81.4	3543	9	BC023950	BC023950 Mus muscu
C 546	36	83.7	254095	14	AC108990	AC108990 Rattus no	C 619	35	81.4	4053	5	GGU65891	U65891 Gallus gall
C 547	36	83.7	254376	14	AC106306	AC106306 Rattus no	C 620	35	81.4	4157	5	GGA401469	AJ401469 Chicken h
C 548	36	83.7	254831	14	AC157027	AC157027 Bos tauru	C 621	35	81.4	4208	8	HSB006824	BX640747 Homo sapi
C 549	36	83.7	254946	14	AC096598	AC096598 Rattus no	C 622	35	81.4	6223	6	AX251874	AX251874 Sequence
C 550	36	83.7	256608	9	AL589699	AL589699 Mouse DNA	C 623	35	81.4	7696	6	AX251153	AX251153 Sequence
C 551	36	83.7	257868	14	AC130923	AC130923 Rattus no	C 624	35	81.4	8714	4	BTU19457	UI9457 Bos taurus
C 552	36	83.7	258592	14	AC121750	AC121750 Rattus no	C 625	35	81.4	10006	6	AX344938	AX344938 Sequence
C 553	36	83.7	262277	14	AC107189	AC107189 Rattus no	C 626	35	81.4	16977	5	TFU53213	US3213 Tetraodon f
C 554	36	83.7	262597	14	AC094133	AC094133 Rattus no	C 627	35	81.4	20650	8	HSB36312	Z80998 Human DNA s
C 555	36	83.7	263236	14	AC094070	AC094070 Rattus no	C 628	35	81.4	33347	3	AY316120	AY316120 Unculture
C 556	36	83.7	263315	14	AC161680	AC161680 Bos tauru	C 629	35	81.4	34335	8	AC000094	AC000094 Homo sapi
C 557	36	83.7	263341	14	AC094581	AC094581 Rattus no	C 630	35	81.4	40442	8	AC144863	AC144863 Homo sapi
C 558	36	83.7	263701	14	AC094442	AC094442 Rattus no	C 631	35	81.4	40574	14	AC087352	AC087352 Homo sapi
C 559	36	83.7	264110	14	AC122626	AC122626 Rattus no	C 632	35	81.4	40965	14	AC012633	AC012633 Homo sapi
C 560	36	83.7	265307	14	AC095842	AC095842 Rattus no	C 633	35	81.4	43900	8	AC005265	AC005265 Homo sapi
C 561	36	83.7	267509	2	CNS073GH	AL590450 chromosom	C 634	35	81.4	43919	8	AL446023	AL446023 Human DNA
C 562	36	83.7	268324	5	AL954838	AL954838 Zebrafieh	C 635	35	81.4	44005	8	AC000081	AC000081 Homo sapi
C 563	36	83.7	271832	5	AV739096	AV739096 Takifugu	C 636	35	81.4	46387	14	U82212	U82212 Homo sapien
C 564	36	83.7	271887	14	AC161670	AC161670 Bos tauru	C 637	35	81.4	47417	14	AC100702	AC100702 Mus muscu
C 565	36	83.7	277782	14	AC105860	AC105860 Rattus no	C 638	35	81.4	50191	8	AL450471	AL450471 Human DNA
C 566	36	83.7	284763	5	CR759869	CR759869 Zebrafieh	C 639	35	81.4	51878	8	AC112233	AC112233 Homo sapi
C 567	36	83.7	293266	14	AC094744	AC094744 Rattus no	C 640	35	81.4	57032	14	AC100694	AC100694 Mus muscu
C 568	36	83.7	319087	14	AC135770	AC135770 Rattus no	C 641	35	81.4	58441	14	AC100031	AC100031 Mus muscu
C 569	36	83.7	329931	14	AC123088	AC123088 Rattus no	C 642	35	81.4	58437	8	BX276094	BX276094 Human DNA
C 570	36	83.7	337658	9	AC121114	AC121114 Mus muscu	C 643	35	81.4	58500	14	AC166515	AC166515 Bos tauru
C 571	36	83.7	341520	2	AE003498	AB003498 Drosophi	C 644	35	81.4	59121	2	U41109	U41109 Caenorhabdi
C 572	36	83.7	343714	14	AC140832	AC140832 Homo sapi	C 645	35	81.4	60927	9	AL732491	AL732491 Mouse DNA
C 573	36	83.7	345012	1	BX572607	BX572607 Rhodosphe	C 646	35	81.4	63954	14	AC087789	AC087789 Homo sapi
C 574	35	81.4	349980	6	AX573238	AX573238 Sequence	C 647	35	81.4	68784	5	AY277971	AY277971 Takifugu
C 575	35	81.4	138	15	AF317966	AF317966 Arabidops	C 648	35	81.4	68986	15	T22111	AC012190 Arabidops
C 576	35	81.4	199	6	BD040203	BD040203 Sequence	C 649	35	81.4	70017	14	AC166702	AC166702 Bos tauru
C 577	35	81.4	264	15	AX904670	AX904670 Sequence	C 650	35	81.4	70653	14	AC124622	AC124622 Mus tauru
C 578	35	81.4	564	10	AV021665	AV021665 Oryza sat	C 651	35	81.4	70741	15	BX842615	BX842615 Neurospor
C 579	35	81.4	573	10	BV277583	BV277583 S232P637F	C 652	35	81.4	72533	14	AC013619	AC013619 Homo sapi
C 580	35	81.4	581	10	G94127	G94127 S208P6525FH	C 653	35	81.4	75892	14	AC156260	AC156260 Medicago
C 581	35	81.4	601	6	AR663166	AR663166 Sequence	C 654	35	81.4	77134	8	AC146380	AC146380 Pan trogl
C 582	35	81.4	614	10	BV337596	BV337596 S230P6508	C 655	35	81.4	78336	8	HS355N11	AL031965 Human DNA
C 583	35	81.4	639	10	BV050791	BV050791 S212P620	C 656	35	81.4	78824	14	AC164860	AC164860 Bos tauru
C 584	35	81.4	674	10	BV258923	BV258923 S235P6107	C 657	35	81.4	79363	14	BX294172	Continuation (5 of
C 585	35	81.4	684	6	AX412600	AX412600 Sequence	C 658	35	81.4	80052	14	AC165528	AC165528 Bos tauru
C 586	35	81.4	684	6	AX506268	AX506268 Sequence	C 659	35	81.4	80466	14	AC027822	AC027822 Homo sapi
C 587	35	81.4	684	6	AX651867	AX651867 Sequence	C 660	35	81.4	81184	8	AC006144	AC006144 Homo sapi

C 661	35	81.4	81831	8	AL365506 Human DNA	C 734	35	81.4	118153	8	HS249C1	AL022154 Human DNA
C 662	35	81.4	82464	5	BX547932 Zebrafish	C 735	35	81.4	121112	15	AC144482	AC144482 Medicago
C 663	35	81.4	83811	5	CR847883 Zebrafish	C 736	35	81.4	122493	15	AC023270	AC023270 Homo sapi
C 664	35	81.4	83881	8	AC113609 Homo sapi	C 737	35	81.4	122972	14	AL691434	AL691434 Homo sapi
C 665	35	81.4	83989	15	H0505A02	C 738	35	81.4	123768	14	BX323885	BX323885 Homo sapi
C 666	35	81.4	84016	8	AL591375 Human DNA	C 739	35	81.4	125311	8	AC068758	AC068758 Homo sapi
C 667	35	81.4	85592	14	AL512630 Mus muscu	C 740	35	81.4	127089	15	AP004165	AP004165 Oryza sat
C 668	35	81.4	85638	8	AC108126 Homo sapi	C 741	35	81.4	127341	14	AC099527	AC099527 Felis cat
C 669	35	81.4	85992	15	ATP2G14 Arabidops	C 742	35	81.4	127347	14	AC166479	AC166479 Bos tauru
C 670	35	81.4	88110	8	AC108470 Homo sapi	C 743	35	81.4	128920	14	AC128886	AC128886 Rattus no
C 671	35	81.4	88824	8	AP001344 Homo sapi	C 744	35	81.4	129010	8	AL589666	AL589666 Human DNA
C 672	35	81.4	89039	5	BX548062 Zebrafish	C 745	35	81.4	131402	9	AL773583	AL773583 Mouse DNA
C 673	35	81.4	89153	8	AC026714 Homo sapi	C 746	35	81.4	132097	14	AC131378	AC131378 Strongylo
C 674	35	81.4	89328	6	AX329823 Sequence	C 747	35	81.4	132614	5	AP002014	AP002014 Homo sapi
C 675	35	81.4	89328	8	H8398C22 Zebrafish	C 748	35	81.4	133503	8	BX649425	BX649425 Zebrafish
C 676	35	81.4	93212	15	AC005936 Arabidops	C 749	35	81.4	133943	5	CR382326	CR382326 Zebrafish
C 677	35	81.4	93212	15	AC137999 Oryza sat	C 750	35	81.4	133974	14	AC108894	AC108894 Bos tauru
C 678	35	81.4	97152	15	AC164796 Bos tauru	C 751	35	81.4	134825	15	AC099325	AC099325 Oryza sat
C 679	35	81.4	101327	5	BX530015 Zebrafish	C 752	35	81.4	134845	8	HS863X19	Z92547 Human DNA s
C 680	35	81.4	101546	14	AP004004 Oryza sat	C 753	35	81.4	134979	8	AL359963	AL359963 Human DNA
C 681	35	81.4	102387	8	AC022114 Homo sapi	C 754	35	81.4	135310	14	AC146087	AC146087 Pan trogl
C 682	35	81.4	102842	9	AL671903 Mouse DNA	C 755	35	81.4	135405	8	AC000025	AC000025 Homo sapi
C 683	35	81.4	103647	15	CNS08CD1 Oryza sat	C 756	35	81.4	135513	8	BS000023	BS000023 Pan trogl
C 684	35	81.4	103966	15	AC125473 Medicago	C 757	35	81.4	135940	15	AC146561	AC146561 Medicago
C 685	35	81.4	104388	14	AC152446 Oryza sat	C 758	35	81.4	135997	14	AC105261	AC105261 Oryza sat
C 686	35	81.4	104835	8	AF130418 Homo sapi	C 759	35	81.4	136116	5	BX470183	BX470183 Zebrafish
C 687	35	81.4	105033	8	AC108937 Homo sapi	C 760	35	81.4	137029	15	AC099040	AC099040 Oryza sat
C 688	35	81.4	106383	8	AC093753 Homo sapi	C 761	35	81.4	137248	14	AC152042	AC152042 Dasytus n
C 689	35	81.4	106979	8	AC083804 Homo sapi	C 762	35	81.4	138799	8	AC145334	AC145334 Pan trogl
C 690	35	81.4	107785	14	BX005451.3 Continuation (4 of	C 763	35	81.4	139632	8	AC145334	AC145334 Pan trogl
C 691	35	81.4	108801	14	CT009618 Danio rer	C 764	35	81.4	139687	8	AL590993	AL590993 Human DNA
C 692	35	81.4	109472	9	AC010001 Mus muscu	C 765	35	81.4	140156	14	AC027261	AC027261 Homo sapi
C 693	35	81.4	110000	1	CR626927.28 Continuation (29 o	C 766	35	81.4	140547	14	AC015597	AC015597 Homo sapi
C 694	35	81.4	110000	1	AP006841.27 Continuation (28 o	C 767	35	81.4	140948	14	AC091795	AC091795 Felis cat
C 695	35	81.4	110000	1	BA000016.06 Continuation (7 of	C 768	35	81.4	141756	8	AC024993	AC024993 Homo sapi
C 696	35	81.4	110000	1	BA000016.22 Continuation (23 o	C 769	35	81.4	143013	5	AL954699	AL954699 Zebrafish
C 697	35	81.4	110000	1	BA000028.09 Continuation (10 o	C 770	35	81.4	143035	8	AC010082	AC010082 Homo sapi
C 698	35	81.4	110000	1	BA000028.16 Continuation (17 o	C 771	35	81.4	143960	8	AC021613	AC021613 Homo sapi
C 699	35	81.4	110000	1	BA000032.00 Continuation (2 of	C 772	35	81.4	143969	15	AP002839	AP002839 Oryza sat
C 700	35	81.4	110000	1	BA000037.22 Continuation (23 o	C 773	35	81.4	144679	14	AC149024	AC149024 Rhinoloph
C 701	35	81.4	110000	1	CP000010.02 Continuation (3 of	C 774	35	81.4	144834	8	AC073345	AC073345 Homo sapi
C 702	35	81.4	110000	1	CP000083.33 Continuation (34 o	C 775	35	81.4	145061	15	AC160012	AC160012 Medicago
C 703	35	81.4	110000	6	RA397408.1 Continuation (2 of	C 776	35	81.4	145137	9	AC115935	AC115935 Mus muscu
C 704	35	81.4	110000	14	AC098001.1 Continuation (2 of	C 777	35	81.4	145430	8	AC073606	AC073606 Homo sapi
C 705	35	81.4	110000	14	AC095084.1 Continuation (2 of	C 778	35	81.4	146575	14	AC144524	AC144524 Homo sapi
C 706	35	81.4	110000	14	AC105976.1 Continuation (2 of	C 779	35	81.4	146833	8	AC010998	AC010998 Homo sapi
C 707	35	81.4	110000	14	AC115960.1 Continuation (2 of	C 780	35	81.4	147038	8	AL133326	AL133326 Human DNA
C 708	35	81.4	110000	14	AC115960.2 Continuation (3 of	C 781	35	81.4	147156	5	CR394572	CR394572 Zebrafish
C 709	35	81.4	110000	14	AC151717.1 Continuation (2 of	C 782	35	81.4	147274	14	AP005588	AP005588 Oryza sat
C 710	35	81.4	110000	14	AP002753.1 Continuation (2 of	C 783	35	81.4	147305	8	AC008769	AC008769 Homo sapi
C 711	35	81.4	110000	14	BX005451.0 Mus muscu	C 784	35	81.4	148060	15	AP003334	AP003334 Oryza sat
C 712	35	81.4	110000	14	BX936368.2 Continuation (3 of	C 785	35	81.4	148352	14	AC109813	AC109813 Homo sapi
C 713	35	81.4	110000	14	CT005272.24 Continuation (25 o	C 786	35	81.4	148491	8	AC026753	AC026753 Homo sapi
C 714	35	81.4	110000	14	TANW2.12 Continuation (13 o	C 787	35	81.4	148496	14	AC134956	AC134956 Tetraodon
C 715	35	81.4	110000	15	AP008215.042 Continuation (43 o	C 788	35	81.4	148822	14	AC058817	AC058817 Homo sapi
C 716	35	81.4	110000	15	AP008216.070 Continuation (71 o	C 789	35	81.4	148991	8	AL622774	AL622774 Human DNA
C 717	35	81.4	110000	15	AP008218.035 Continuation (36 o	C 790	35	81.4	149308	8	AC005527	AC005527 Homo sapi
C 718	35	81.4	110000	15	AP008207.218 Continuation (219	C 791	35	81.4	149542	15	AP003616	AP003616 Oryza sat
C 719	35	81.4	110000	15	AP008207.251 Continuation (252	C 792	35	81.4	150487	5	BX255900	BX255900 Zebrafish
C 720	35	81.4	110000	15	AP008208.295 Continuation (296	C 793	35	81.4	150594	14	AC108889	AC108889 Bos tauru
C 721	35	81.4	110000	15	AP008209.175 Continuation (176	C 794	35	81.4	150810	8	AC018500	AC018500 Homo sapi
C 722	35	81.4	110000	15	AP008211.163 Continuation (164	C 795	35	81.4	151030	9	AL805936	AL805936 Mouse DNA
C 723	35	81.4	110000	15	AP008212.239 Continuation (240	C 796	35	81.4	151073	8	AC073141	AC073141 Homo sapi
C 724	35	81.4	110000	15	AP008212.240 Continuation (241	C 797	35	81.4	151262	9	AL928698	AL928698 Mouse DNA
C 725	35	81.4	110000	15	AP008212.303 Continuation (304	C 798	35	81.4	151263	15	AP004333	AP004333 Oryza sat
C 726	35	81.4	110000	15	AP008213.294 Continuation (295	C 799	35	81.4	151700	8	AC133528	AC133528 Homo sapi
C 727	35	81.4	110992	8	AC078805 Homo sapi	C 800	35	81.4	152489	9	AC139576	AC139576 Mus muscu
C 728	35	81.4	113200	15	AC126014 Medicago	C 801	35	81.4	153089	8	AC146510	AC146510 Pan trogl
C 729	35	81.4	114143	15	AC127170 Medicago	C 802	35	81.4	153193	14	AC158792	AC158792 Homo sapi
C 730	35	81.4	114452	15	AC148289 Medicago	C 803	35	81.4	154199	14	AC013639	AC013639 Homo sapi
C 731	35	81.4	115134	15	AC126786 Medicago	C 804	35	81.4	154361	14	AC084189	AC084189 Homo sapi
C 732	35	81.4	115857	15	AC147407 Medicago	C 805	35	81.4	155295	14	AC055859	AC055859 Homo sapi
C 733	35	81.4	115857	15	AC147407 Medicago	C 806	35	81.4	155382	8	AC004911	AC004911 Homo sapi

807	35	81.4	155723	8	AP002847	AP002847 Homo sapi	880	35	81.4	168665	14	AL591965	AL591965 Homo sapi
c 808	35	81.4	155889	9	AL591411	AL591411 Mouse DNA	c 881	35	81.4	168853	14	AC055117	AC055117 Homo sapi
c 809	35	81.4	156471	14	AC013568	AC013568 Homo sapi	c 882	35	81.4	169565	14	AC079108	AC079108 Homo sapi
810	35	81.4	156504	14	AC148250	AC148250 Ootlemur	c 883	35	81.4	169604	14	AC068849	AC068849 Homo sapi
c 811	35	81.4	156727	14	CR936976	CR936976 Homo sapi	c 884	35	81.4	170126	14	AC149177	AC149177 Papio anu
c 812	35	81.4	156951	8	AC011146	AC011146 Homo sapi	c 885	35	81.4	170295	5	BX663508	BX663508 Zebrafish
c 813	35	81.4	157452	14	AC020841	AC020841 Mus muscu	c 886	35	81.4	170410	14	AC152019	AC152019 Papio anu
c 814	35	81.4	158027	9	AC125093	AC125093 Mus muscu	c 887	35	81.4	170510	15	AP005589	AP005589 Oryza sat
c 815	35	81.4	158135	9	AC125093	AC125093 Mus muscu	c 888	35	81.4	171482	8	AC034247	AC034247 Homo sapi
c 816	35	81.4	158321	15	AP003626	AP003626 Oryza sat	c 889	35	81.4	171732	14	AC021994	AC021994 Homo sapi
c 817	35	81.4	158347	9	AL133401	AL133401 Mouse DNA	c 890	35	81.4	171757	8	CNS01D80	AL121857 Human chr
c 818	35	81.4	158357	8	CNS078FC	AL152414 Human chr	c 891	35	81.4	172154	14	AC147733	AC147733 Macropus
c 819	35	81.4	158882	14	AL732407	AL732407 Homo sapi	c 892	35	81.4	172476	8	AC007736	AC007736 Homo sapi
c 820	35	81.4	159385	14	AC024708	AC024708 Homo sapi	c 893	35	81.4	172678	8	AC0078916	AC0078916 Homo sapi
c 821	35	81.4	159473	9	AC1555167	AC155167 Mus muscu	c 894	35	81.4	172716	14	AC115383	AC115383 Rattus no
c 822	35	81.4	159506	8	HS3418	AL021918 Human DNA	c 895	35	81.4	172816	8	AC093899	AC093899 Homo sapi
c 823	35	81.4	159610	14	AC012137	AC012137 Homo sapi	c 896	35	81.4	173093	14	AC156079	AC156079 Bos tauru
c 824	35	81.4	159620	8	AL354835	AL354835 Human DNA	c 897	35	81.4	173861	14	AC141417	AC141417 Papio anu
c 825	35	81.4	160159	14	AC126896	AC126896 Rattus no	c 898	35	81.4	173874	14	AC068011	AC068011 Homo sapi
c 826	35	81.4	160590	14	AC127392	AC127392 Rattus no	c 899	35	81.4	174143	14	AC113806	AC113806 Rattus no
c 827	35	81.4	160855	8	AC109925	AC109925 Homo sapi	c 900	35	81.4	174351	9	AC103351	AC103351 Mus muscu
c 828	35	81.4	160912	8	AP003471	AP003471 Homo sapi	c 901	35	81.4	174507	9	AC132955	AC132955 Mus muscu
c 829	35	81.4	161040	14	AC021912	AC021912 Homo sapi	c 902	35	81.4	174545	8	AC103801	AC103801 Homo sapi
c 830	35	81.4	161193	14	AC011033	AC011033 Homo sapi	c 903	35	81.4	174547	8	AP005136	AP005136 Homo sapi
c 831	35	81.4	161528	9	AC102763	AC102763 Mus muscu	c 904	35	81.4	174676	9	AC157606	AC157606 Mus muscu
c 832	35	81.4	161531	14	CR936905	CR936905 Danio rer	c 905	35	81.4	175268	8	AC100775	AC100775 Homo sapi
c 833	35	81.4	161668	14	AC157547	AC157547 Mus muscu	c 906	35	81.4	175360	14	AC018948	AC018948 Homo sapi
c 834	35	81.4	161685	14	AC013629	AC013629 Homo sapi	c 907	35	81.4	176316	8	AC010798	AC010798 Homo sapi
c 835	35	81.4	162064	8	AL390882	AL390882 Human DNA	c 908	35	81.4	176316	8	AC114808	AC114808 Homo sapi
c 836	35	81.4	162082	15	AC144742	AC144742 Oryza sat	c 909	35	81.4	176392	14	AC120903	AC120903 Rattus no
c 837	35	81.4	162334	14	AC138068	AC138068 Pan trogl	c 910	35	81.4	176426	8	AC007370	AC007370 Homo sapi
c 838	35	81.4	162341	14	AC163856	AC163856 Bos tauru	c 911	35	81.4	176570	8	AC009499	AC009499 Homo sapi
c 839	35	81.4	162573	8	CNS0008BL	AL079304 Human chr	c 912	35	81.4	176581	8	AC117535	AC117535 Homo sapi
c 840	35	81.4	163514	14	AC011112	AC011112 Homo sapi	c 913	35	81.4	176668	14	AC027096	AC027096 Homo sapi
c 841	35	81.4	163681	8	AL136992	AL136992 Human DNA	c 914	35	81.4	176932	8	AC016772	AC016772 Homo sapi
c 842	35	81.4	163795	8	CNS059EU	AL359396 Human chr	c 915	35	81.4	176981	9	AC144761	AC144761 Mus muscu
c 843	35	81.4	163796	8	AC093725	AC093725 Homo sapi	c 916	35	81.4	177076	14	AC092949	AC092949 Homo sapi
c 844	35	81.4	163811	14	AC130407	AC130407 Homo sapi	c 917	35	81.4	177127	5	AC026287	AC026287 Homo sapi
c 845	35	81.4	163957	8	AC104786	AC104786 Homo sapi	c 918	35	81.4	177177	5	BX510312	BX510312 Zebrafish
c 846	35	81.4	163957	14	AC151372	AL51372 Actus nan	c 919	35	81.4	177206	5	CR354558	CR354558 Zebrafish
c 847	35	81.4	164002	9	AC153135	AC153135 Mus muscu	c 920	35	81.4	177240	9	AC151283	AC151283 Mus muscu
c 848	35	81.4	164232	14	AC067792	AC067792 Homo sapi	c 921	35	81.4	177420	9	AC105162	AC105162 Mus muscu
c 849	35	81.4	164685	9	AC128667	AC128667 Mus muscu	c 922	35	81.4	177617	5	AL840638	AL840638 Zebrafish
c 850	35	81.4	164724	14	CR555297	CR555297 Danio rer	c 923	35	81.4	177978	8	AC009073	AC009073 Homo sapi
c 851	35	81.4	164774	8	AL512639	AL512639 Human DNA	c 924	35	81.4	177978	8	AC009126	AC009126 Homo sapi
c 852	35	81.4	164888	14	AC024101	AC024101 Homo sapi	c 925	35	81.4	178142	8	AC013546	AC013546 Homo sapi
c 853	35	81.4	165590	8	AL512885	AL512885 Human DNA	c 926	35	81.4	178193	9	AC101702	AC101702 Mus muscu
c 854	35	81.4	165655	8	AL336962	AL336962 Human DNA	c 927	35	81.4	178416	14	BX936329	BX936329 Danio rer
c 855	35	81.4	165878	14	AC025247	AC025247 Homo sapi	c 928	35	81.4	178426	14	AC159831	AC159831 Bos tauru
c 856	35	81.4	165973	8	AC097515	AC097515 Homo sapi	c 929	35	81.4	178560	9	AC134839	AC134839 Mus muscu
c 857	35	81.4	165993	5	AL772362	AL772362 Zebrafish	c 930	35	81.4	178825	14	AP001203	AP001203 Homo sapi
c 858	35	81.4	166114	5	AL807829	AL807829 Zebrafish	c 931	35	81.4	178952	8	AP418272	AP418272 Homo sapi
c 859	35	81.4	166189	14	CR376746	CR376746 Danio rer	c 932	35	81.4	178968	14	CR450709	CR450709 Danio rer
c 860	35	81.4	166228	5	AL929341	AL929341 Zebrafish	c 933	35	81.4	179014	14	AC109984	AC109984 Rattus no
c 861	35	81.4	166625	8	BS000032	BS000032 Pan trogl	c 934	35	81.4	179139	14	AC022484	AC022484 Homo sapi
c 862	35	81.4	166656	14	AC025331	AC025331 Homo sapi	c 935	35	81.4	179170	14	AC165322	AC165322 Mus muscu
c 863	35	81.4	166700	15	AP004685	AP004685 Oryza sat	c 936	35	81.4	179321	9	AC101719	AC101719 Mus muscu
c 864	35	81.4	167178	14	AC025672	AC025672 Homo sapi	c 937	35	81.4	179342	14	AC123285	AC123285 Rattus no
c 865	35	81.4	167237	8	AC007938	AC007938 Homo sapi	c 938	35	81.4	179452	8	AC116010	AC116010 Homo sapi
c 866	35	81.4	167288	14	AC139120	AC139120 Didelphis	c 939	35	81.4	179498	5	AL845314	AL845314 Zebrafish
c 867	35	81.4	167319	8	AC090541	AC090541 Homo sapi	c 940	35	81.4	179741	9	AC153497	AC153497 Mus muscu
c 868	35	81.4	167363	9	AC141640	AC141640 Mus muscu	c 941	35	81.4	179904	14	AC149705	AC149705 Bos tauru
c 869	35	81.4	167417	14	AC151166	AC151166 Bos tauru	c 942	35	81.4	179904	6	AR659536	AR659536 Sequence
c 870	35	81.4	167472	14	CR936499	CR936499 Danio rer	c 943	35	81.4	179905	6	AR659664	AR659664 Sequence
c 871	35	81.4	167502	14	AC018909	AC018909 Homo sapi	c 944	35	81.4	180315	9	AC107843	AC107843 Mus muscu
c 872	35	81.4	167927	14	AC162782	AC162782 Danio rer	c 945	35	81.4	180424	8	AC106860	AC106860 Homo sapi
c 873	35	81.4	167935	14	AC015638	AC015638 Homo sapi	c 946	35	81.4	180584	14	AC1010703	AC1010703 Homo sapi
c 874	35	81.4	167961	14	AP002422	AP002422 Homo sapi	c 947	35	81.4	180689	14	AC124913	AC124913 Sus scrof
c 875	35	81.4	168242	14	AC124236	AC124236 Mus muscu	c 948	35	81.4	180692	9	AC134559	AC134559 Mus muscu
c 876	35	81.4	168347	8	AC079756	AC079756 Homo sapi	c 949	35	81.4	180967	14	AC166163	AC166163 Mus muscu
c 877	35	81.4	168443	14	AC137940	AC137940 Mus muscu	c 950	35	81.4	181168	14	AC055888	AC055888 Homo sapi
c 878	35	81.4	168508	14	AC149240	AC149240 Muntiacus	c 951	35	81.4	181237	14	CR394548	CR394548 Danio rer
c 879	35	81.4	168602	14	AF271406	AF271406 Homo sapi	c 952	35	81.4	181271	14	CR749171	CR749171 Danio rer

```
953 35 81.4 181292 14 CR388363
c 954 35 81.4 181307 5 BX649358
955 35 81.4 181322 8 AC027018
c 956 35 81.4 181534 9 AC102553
957 35 81.4 182157 14 AC155759
958 35 81.4 182209 8 AL357060
c 959 35 81.4 182270 9 AC160965
960 35 81.4 182288 8 AC018362
c 961 35 81.4 182311 9 AC114677
c 962 35 81.4 182417 14 AC025291
963 35 81.4 182781 14 AC087814
c 964 35 81.4 182883 9 AC134859
965 35 81.4 183113 14 AC155581
c 966 35 81.4 183227 9 AC154372
967 35 81.4 183338 14 BX897753
c 968 35 81.4 183445 14 AC026212
969 35 81.4 183478 9 AC148324
c 970 35 81.4 183514 14 AC012412
971 35 81.4 183603 14 AC128037
c 972 35 81.4 183761 14 AC158788
973 35 81.4 184043 14 AC090296
c 974 35 81.4 184164 9 AC126944
975 35 81.4 184165 14 AC155037
c 976 35 81.4 184303 9 AL807734
977 35 81.4 184463 9 AC126553
c 978 35 81.4 184510 14 AC018894
979 35 81.4 184626 5 BX470258
c 980 35 81.4 185321 14 AC026153
981 35 81.4 185571 8 AC133781
c 982 35 81.4 185608 8 AC027126
983 35 81.4 185765 14 AC125883
c 984 35 81.4 186074 8 AP005061
985 35 81.4 186134 9 AL591390
c 986 35 81.4 186186 14 AC134752
987 35 81.4 186222 8 AC146229
c 988 35 81.4 186276 8 AC024168
989 35 81.4 186441 9 AC122258
c 990 35 81.4 187043 9 AL591586
991 35 81.4 187658 9 AC083895
c 992 35 81.4 187683 5 BX510336
993 35 81.4 187946 14 AC073961
c 994 35 81.4 188027 14 AC166002
995 35 81.4 188335 9 AC118733
c 996 35 81.4 188518 9 AC126454
997 35 81.4 188789 14 AP001446
c 998 35 81.4 188946 9 AC127236
999 35 81.4 189840 8 AC012320
1000 35 81.4 190256 9 AC121302

ALIGNMENTS

RESULT 1
CQ559387
LOCUS CQ559387 65 bp DNA
DEFINITION Sequence 29022 from Patent WO0210449.
ACCESSION CQ559387
VERSION CQ559387.1 GI:41525814
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
1 Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
AUTHORS Oligonucleotide library for detecting rna transcripts and splice
TITLE variants that populate a transcriptome
JOURNAL Patent: WO 0210449-A 29022 07-FEB-2002;
COMPUGEN INC. (US)
FEATURES Location/Qualifiers
source 1..65
/organism="Mus musculus"

CR388363 Danio rer
BX649358 Zebrafish
AC027018 Homo sapi
AC102553 Mus muscu
AC155759 Bos tauru
AL357060 Human DNA
AC160965 Mus muscu
AC018362 Homo sapi
AC114677 Mus muscu
AC025291 Homo sapi
AC087814 Homo sapi
AC134859 Mus muscu
AC155581 Zea mays
AC154372 Mus muscu
BX897753 Danio rer
AC026212 Homo sapi
AC148324 Mus muscu
AC012412 Homo sapi
AC128037 Rattus no
AC158788 Mus muscu
AC090296 Homo sapi
AC126944 Mus muscu
AC155037 Bos tauru
AL807734 Mouse DNA
AC126553 Mus muscu
AC018894 Homo sapi
BX470258 Zebrafish
AC026153 Homo sapi
AC133781 Homo sapi
AC027126 Homo sapi
AC125883 Rattus no
AP005061 Homo sapi
AL591390 Mouse DNA
AC134752 Rattus no
AC146229 Pan trogl
AC024168 Homo sapi
AC122258 Mus muscu
AL591586 Mouse DNA
AC083895 Mus muscu
BX510336 Zebrafish
AC073961 Homo sapi
AC166002 Nomauscu
AC118733 Mus muscu
AC126454 Mus muscu
AP001446 Homo sapi
AC127236 Mus muscu
AC012320 Homo sapi
AC121302 Mus muscu

/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores: 2.55 Length: 65
Pred. No.: 43.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-7 (1-9) x CQ559387 (1-65)
Qy 1 SerLeuGlnThrSerTyxValPheLeu 9
Db 29 TCCCTGCAGACTTCCTATGCTTCCTA 55

RESULT 2
CQ687716 290 bp DNA linear PAT 03-FEB-2004
LOCUS CQ687716
DEFINITION Sequence 32642 from Patent WO02070737.
ACCESSION CQ687716
VERSION CQ687716.1 GI:42218962
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
1 Liew,C.C., Marshall,W.E. and Zhang,H.
AUTHORS Compositions and methods relating to osteoarthritis
TITLE Patent: WO 02070737-A 32642 12-SEP-2002;
JOURNAL Chondrogene Inc. (CA)
FEATURES Location/Qualifiers
source 1..290
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 9.02 Length: 290
Pred. No.: 43.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-7 (1-9) x CQ687716 (1-290)
Qy 1 SerLeuGlnThrSerTyxValPheLeu 9
Db 32 TCCCTGCAGAACTCTATGCTTCCTG 58

RESULT 3
CQ920916 475 bp DNA linear PAT 23-NOV-2004
LOCUS CQ920916
DEFINITION Sequence 2116 from Patent WO2004097052.
ACCESSION CQ920916
VERSION CQ920916.1 GI:56210857
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE Hominidae; Homo.
JOURNAL Burczynski,M.B., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
Strahs,A., Immerman,P. and Dörner,A.J.
METHODS Patent: WO 2004097052-A 2116 11-NOV-2004;
FEATURES Location/Qualifiers
source 1..65
/organism="Mus musculus"
```

source	location/coordinates
1. .901	/organism="Canis sp."

```
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN
Alignment Scores:
Pred. No.: 23.5 Length: 901
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX316088 (1-901)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 608 TCCCTGCAGACTTCTTATGTTCTCTA 634

RESULT 7
AX829164
LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 Mack,D.H., Gish,K.C. and Afar,D.
AUTHORS Methods of diagnosis of breast cancer, compositions and methods of
TITLE screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)
FEATURES
source
1. 927
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 24.1 Length: 927
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX829164 (1-927)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 709 TCCCTGCAGACTTCTTATGTTCTCTG 735

RESULT 8
AX467373
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE Felis sp.
ORGANISM Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 Myers,K., Drury,N. and Carroll,M.
AUTHORS Polypeptide
TITLE Patent: WO 0238612-A 3 16-MAY-2002;
JOURNAL
```

```

Oxford Biomedica (UK) Limited (GB)
Location/Qualifiers
1. 1260
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"

ORIGIN
Alignment Scores:
Pred. No.: 31.2 Length: 1260
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX467373 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1048 TCCCTGCAGACTTCTTATGTTCTTCTA 1074

RESULT 9
AX821533
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE Felis catus (cat)
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 Carroll,M.M., Kingman,S.M. and Redchenko,I.M.
AUTHORS MHC class I peptide epitopes from the human 5t4 tumor-associated
TITLE antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. 1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN
Alignment Scores:
Pred. No.: 31.2 Length: 1260
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX821533 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1048 TCCCTGCAGACTTCTTATGTTCTTCTA 1074

RESULT 10
AX821548
LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE Felis catus (cat)
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
```

REFERENCE
1
AUTHORS Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
TITLE MHC class II peptide epitope of 5t4 antigen
JOURNAL Patent: WO 0368815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN
Alignment Scores:
Pred. No.: 31.2 Length: 1260
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX821548 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1048 TCCCTGCAGACTTCTATGCTTTCTA 1074

RESULT 11
BD249731
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
1 (bases 1 to 1263)
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PR 18-NOV-1999 JP 2000582415
PR 30-JUL-1999 GB 9917995.4

PI MILES WILLIAM CARROLL,KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
PC C07K19/00,
PC C12N15/00
CC Polypeptide
FT Key Location/Qualifiers
FT source 1..1263
/organism="Homo sapiens (human)"
FT Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
Alignment Scores:
Pred. No.: 31.3 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x BD249731 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCAAACTCTATGCTTCCG 1077

RESULT 12
AX025011
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO0029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 31.3 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX025011 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCAAACTCTATGCTTCCG 1077

RESULT 13
AX149553
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
Ellard,P.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="5T4"

ORIGIN
Alignment Scores:
Pred. No.: 31.3 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX025011 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCAAACTCTATGCTTCCG 1077

RESULT 13
AX149553
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
Ellard,P.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="5T4"

ORIGIN
Alignment Scores:
Pred. No.: 31.3 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

US-10-774-176-7 (1-9) x BD249731 (1-1263)

Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX149553 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 1051 TCCCTGCAGACTTCTATGTCTTCCTA 1077

RESULT 14
 AX316086
 LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
 DEFINITION Sequence 1 from Patent EP1160323.
 ACCESSION AX316086
 VERSION AX316086.1 GI:17899278
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Carroll, M.W. and Myers, K.A.
 TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
 JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1..1263
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 31.3 Length: 1263
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX316086 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 1051 TCCCTGCAGACTTCTATGTCTTCCTG 1077

RESULT 15
 AX467371
 LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
 DEFINITION Sequence 1 from Patent WO0238612.
 ACCESSION AX467371
 VERSION AX467371.1 GI:21900602
 KEYWORDS
 SOURCE Canis sp.
 ORGANISM Canis sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.

REFERENCE
 AUTHORS Myers, K., Drury, N. and Carroll, M.
 TITLE Polypeptide
 JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1..1263
 /organism="Canis sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:9616"

ORIGIN

Alignment Scores:
 Pred. No.: 31.3 Length: 1263

Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX467371 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 1051 TCCCTGCAGACTTCTATGTCTTCCTA 1077

RESULT 16
 BD249732
 LOCUS BD249732 1281 bp DNA linear PAT 17-JUL-2003
 DEFINITION Polypeptide.
 ACCESSION BD249732
 VERSION BD249732.1 GI:33059502
 KEYWORDS JP 2002530060-A/2.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Carroll, M.W. and Myers, K.A.
 TITLE Polypeptide
 JOURNAL Patent: JP 2002530060-A 2 17-SEP-2002;
 OXFORD BIOMEDICA LTD

COMMENT
 OS Mus musculus (mouse)
 PN JP 2002530060-A/2
 PD 17-SEP-2002
 PP 18-NOV-1999 JP 2000582415
 PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
 30-JUL-1999 GB 9917995.4
 PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
 PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K14/06, C07K14/065,
 C07K19/00,
 PC C12N15/00,
 CC Polypeptide
 FH Key
 FT source
 1..1281
 Location/Qualifiers
 /organism="Mus musculus (mouse)"

FEATURES
 source
 1..1281
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"

ORIGIN

Alignment Scores:
 Pred. No.: 31.7 Length: 1281
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x BD249732 (1-1281)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 1069 TCCCTGCAGACTTCTATGTCTTCCTA 1095

RESULT 17
 AX025012
 LOCUS AX025012 1281 bp DNA linear PAT 15-SEP-2000
 DEFINITION Sequence 2 from Patent WO029428.
 ACCESSION AX025012
 VERSION AX025012.1 GI:10184933
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 Carroll, M.W. and Myers, K.A.

AUTHORS Polypeptide

TITLE Patent: WO 0029428-A 2 25-MAY-2000;

JOURNAL CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)

FEATURES Location/Qualifiers

source
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 31.7 Length: 1281
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX025012 (1-1281)

QY 1 SerLeuGlnThrSerTyValPheLeu 9

Db 1069 TCCCTGCAGACTTCCTATGTCCTTA 1095

RESULT 18

AX316087 LOCUS AX316087 1281 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 2 from Patent EP1160323.

ACCESSION AX316087

VERSION AX316087.1 GI:17899279

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE 1 Carroll, M.W. and Myers, K.A.

AUTHORS 5t4 tumour-associated antigen for use in tumour immunotherapy

TITLE Patent: EP 1160323-A 2 05-DEC-2001;

JOURNAL Oxford Biomedica (UK) Limited (GB)

FEATURES Location/Qualifiers

source
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 31.7 Length: 1281
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX316087 (1-1281)

QY 1 SerLeuGlnThrSerTyValPheLeu 9

Db 1069 TCCCTGCAGACTTCCTATGTCCTTA 1095

RESULT 19

AX316087 LOCUS AX316087 1281 bp DNA linear PRI 18-APR-2005
DEFINITION Homo sapiens 5t4 gene for 5t4 oncofoetal antigen.

ACCESSION Z29083

VERSION Z29083.1 GI:435654

KEYWORDS 5t4 gene; 5t4 oncofoetal antigen.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 2053)

AUTHORS Myers, K.A., Rahi-Saund, V., Davison, M.D., Young, J.A., Cheater, A.J.
and Stern, P.L.

TITLE Isolation of a cDNA encoding 5t4 oncofoetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats

JOURNAL J. Biol. Chem. 269 (12), 9319-9324 (1994)

PUBMED 8132670

REFERENCE 2 (bases 1 to 2053)

AUTHORS Myers, K.A.

TITLE Direct Submission
Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK

FEATURES

source
1..2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label="N-flank"
85..1347
/evidence="experimental"
/product="5t4 oncofoetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"

misc_RNA

CDS

/translation="MPGCGSRGPAAGDGRRLRLALVLGWSSSPTSSASSFSSS
APFLASVSAQPLPDQPCALCESEARATVCKVNRNLTEVPTDLPAYVRNLFTGNQ
LAVLPAGAFARPPPLAALNLGSLRDEVRAGAFELPSRLQLDLSHNLADLPF
AFSGSNASVSAQPLVELLNHIVPEDSRONRSPGVVAALLAGRALQGLRLELA
SNHLYLPRDVLQAQLPSLRHLDLNNSLSVLTYSFERNLTLESLELDNALKVLHG
TLARLQGLPHIRVFLDNPNWCCDMADMTWLKETEYVQGGDRLTCAIPEKQNRVL
LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGALFLVLVLRNKGKIKKMMHNRDAC
RDHMEGYHYRYEINADPRLTNLSNSDV"
130..171
373..966
/product="Leucine rich repeat region"
/label="LRRS"
966..1119
/product="LRR C-terminal flank"
/label="C-flank"
1153..1215
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

ORIGIN

Alignment Scores:

Pred. No.: 47.2 Length: 2053
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x HS5T40A (1-2053)


```

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1135 TCCCTGCAAACTCTTATGTCCTCTG 1161

RESULT 20
CR855786 2183 bp mRNA linear VRT 03-NOV-2004
LOCUS Xenopus tropicalis finished cDNA, clone TGA020h08.
DEFINITION
ACCESSION CR855786
VERSION CR855786.1 GI:55295318
KEYWORDS
SOURCE
ORGANISM Xenopus tropicalis (Silurana tropicalis)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE
1 (bases 1 to 2183)
AUTHORS Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,
Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,
McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,
Voigt,J. and Zorn,A.M.
DIRECT SUBMISSION
TITLE Xenopus tropicalis
JOURNAL Submitted (03-NOV-2004) Sanger Institute, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
COMMENT Sanger Xenopus tropicalis EST/cDNA project.
This sequence is from a Xenopus Gene Collection (XGC) library, from
a library constructed by Aaron M. Zorn. cDNA was prepared from RNA
extracted from gastrula embryos. EcoRI-NotI cut cDNA was then
ligated into pCG107 with EcoRI at the 5' end and NotI at the 3'
end.
Vector: pCG107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.
FEATURES
source
location/Qualifiers
1..2183
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone_lib="XGC-gastrula"
/dev_stage="gastrula (stage 10.5-13 mixed)"

ORIGIN
Alignment Scores:
Pred. No.: 49.7 Length: 2183
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-7 (1-9) x CR855786 (1-2183)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 726 AGCTTGCAAACTCTTATGTCCTCTG 752

RESULT 21
AF063939 2333 bp mRNA linear ROD 01-JAN-2000
LOCUS Rattus norvegicus 574 oncofetal antigen homolog (574) mRNA,
complete cds.
DEFINITION
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 2333)
AUTHORS Ninkina,N.N. and Buchman,V.L.
TITLE Structure and expression of the rat 574 gene

Unpublished
2 (bases 1 to 2333)
AUTHORS Buchman,V.L.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK
FEATURES
source
location/Qualifiers
1..2333
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="Cerebellum"
/dev_stage="newborn"
1..2333
/gene="574"
1..363
/gene="574"
364..1644
/gene="574"
/codon_start=1
/product="574 oncofetal antigen homolog"
/protein_id="AAF21770.1"
/db_xref="GI:6650212"
/translation="MPGAGSGPSAGDGRRLRLRLALVLLGWVSASAPSSSLPSSSTS
PAAFIAGSAQPPPAERCPAAECSEAAATVKVNRNLLLEVPADLPVYVNNPLTNGQ
MTVLPAGAPARQPLADLAVNLNCSHLKEVGAGAFELHPLGLRLDLSHNPLTNSAF
TFAGNSVSVSTPPLLELILNHIVPPDQRQNGSPFGMVAFGMAALRSGIALRGL
HHLELASHFHLYLPRLDLQPLSKHLDRNNLSVLTYSFASFNLTLESLEDNAL
KVLNKSITLAEQWGLAHVRVFLDNNPWCDCYNMADVSWLKETEVPVDFKARLTCAFPPEK
MNRGLDLTSDLDLDCDQATLPQSLOTSVYFGLIVLALIGAFLLVLYLNKGIKKMMH
NIRDACRDHMSGYHYRYEINADPSLTNLSNSDV"
1645..2333
/gene="574"
2315..2320
/gene="574"

3'UTR
polya_signal

ORIGIN
Alignment Scores:
Pred. No.: 52.5 Length: 2333
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-7 (1-9) x AF063939 (1-2333)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1432 TCCCTGCAAACTCTTATGTCCTCTG 1458

RESULT 22
BD127282 2359 bp DNA linear PAT 18-SEP-2002
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION
ACCESSION BD127282
VERSION BD127282.1 GI:23222227
KEYWORDS JP 2002017375-A/2713.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isegai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2713 22-JAN-2002;
HELIIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/2713

```

PD 22-JAN-2002
PP 07-JUL-2000 JP 2000253172
PI TOSHIO OSA, TETSUO NISHIKAWA, TAKAO ISOgai, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
Location/Qualifiers
FT CDS (424)..(1572).

FEATURES

source
1..2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 53 Length: 2359
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x BD127282 (1-2359)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1474 TCCTGCAAACTCTTATGTCCTCTG 1500

RESULT 23

CQ782724
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

QY782724
Sequence 2864 from Patent EP1396543.
CQ782724
CQ782724.1 GI:45502667
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
Ota, T., Nishikawa, T., Isogai, T., Hayaashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
Primers for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 2864 10-MAR-2004; (JP)
Research Association for Biotechnology (JP)
Location/Qualifiers
1..2359
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAP85958.1"
/db_xref="GI:45502668"

CDS

translation="MPGCGRGPAAAGDRLRLARLALVLGWSSSPSSASSPSSSS
APFLASVAQPLDQCCEAARTRVCRNLTEVDTDFPAYRNLFTGNQ
LAVLPAGAPRPPPLAEALNLHVPEDERQNRSPFGMVVAALLAGRALQGLRLLELA
AFSGSNASVAPSPVLVNLHVPEDERQNRSPFGMVVAALLAGRALQGLRLLELA
SNHLYLPRDVLVLAQLPSLRHLDSNLSVLTYSFRNLTHLSLHLELNALKVHLNG
TLAEQGLPHIRFVLDNPNWCDCHADMTWLKETEYVQGGKRLTCATPEKRNKRLV
LELNSADLDCDPLPPSLQTSYVFLGIVLALIGALFLLVLYLNRKGIKK"

ORIGIN

US-10-774-176-7 (1-9) x AK074786 (1-2359)

Alignment Scores:
Pred. No.: 53 Length: 2359
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x CQ782724 (1-2359)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1474 TCCTGCAAACTCTTATGTCCTCTG 1500

RESULT 24

AK074786

LOCUS

DEFINITION
Homo sapiens CDNA FLJ90305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 574 oncofetal trophoblast glycoprotein gene.

ACCESSION
AK074786

VERSION
AK074786.1 GI:22760460

KEYWORDS
oligo capping; fis (full insert sequence).

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1

AUTHORS

Isogai, T., Ota, T., Nishikawa, T., Hayaashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahari, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuma, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.

NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 2359)

REFERENCE

Isogai, T. and Otsuki, T.

Direct Submission

TITLE

Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

Location/Qualifiers

source

1..2359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 53 Length: 2359
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

QY 1 SerIeuGlnThrSertYrValPheLeu 9
Db 1474 TCCTCGCAACCTCTTATGCTTCCTG 1500

RESULT 25
BD127283
LOCUS 2361 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127283
VERSION BD127283.1 GI:232222228
KEYWORDS JP 2002017375-A/2714.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2361)
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HeliX Research Institute
COMMENT Patent: JP 2002017375-A 2714 22-JAN-2002;
HELI X RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2714
PD 22-JAN-2002
PF 07-JUL-2000 JP 200253172
PI TOSHIO Ota, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINTCHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12P1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
PC Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
(426)..(1685).
FEATURES
source
1. .2361
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 53.1 Length: 2361
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-7 (1-9) x CQ782726 (1-2361)

QY 1 SerIeuGlnThrSertYrValPheLeu 9
Db 1476 TCCTCGCAACCTCTTATGCTTCCTG 1502

RESULT 26
AX961916
LOCUS 2361 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 127 from Patent WO03104277.
ACCESSION AX961916
VERSION AX961916.1 GI:40881326
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1
Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
AUTHORS Stat6 activation gene
TITLE Patent: WO 03104277-A 127 18-DEC-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. .2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
426..1688
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06467.1"
/db_xref="GI:40881327"

QY 1 SerIeuGlnThrSertYrValPheLeu 9
Db 1476 TCCTCGCAACCTCTTATGCTTCCTG 1502

RESULT 26
CQ782726
LOCUS 2361 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2866 from Patent EP1396543.
ACCESSION CQ782726
VERSION CQ782726.1 GI:45502669
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1

LELNSADLDGCDPILPSPSQTSYVFLGIVLALIGAIFLLVLYLNRKGIKKMMHNRDACC
RDHMEGYHYREINADPRLTNLSSNDV"

ORIGIN

Alignment Scores:
Pred. No.: 53.1 Length: 2361
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-7 (1-9) x AX961916 (1-2361)

Qy 1 SerLeuGlnThrSerTyValPheLeu 9

Db 1476 TCCCTGCAAACTCTTATGCTTCCTG 1502

RESULT 28

AK074790

LOCUS AK074790 2361 bp mRNA linear PRI 09-JUL-2005
DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens 574 oncofetal trophoblast glycoprotein gene.

ACCESSION

AK074790

VERSION AK074790.1 GI:22760466

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

AUTHORS

1 Otsuki,T., Ota,T., Nishikawa,T., Hayashi,K., Suzuki,Y.,
Yamamoto,J., Wakamatsu,A., Kimura,K., Sakamoto,K., Hatanou,N.,
Kawai,T., Ishii,S., Saito,K., Kojima,S., Sugiyama,T., Ono,T.,
Okano,K., Yoshikawa,Y., Aotseuka,S., Sasaki,N., Hattori,A.,
Okumura,K., Nagai,K., Sugano,S. and Isogai,T.

Signal Sequence and Keyword Trap in silico for Selection of
Full-length Human cDNAs Encoding Secretion or Membrane Proteins
from Oligo-Capped cDNA Libraries
DNA Res. 12, 117-126 (2005)

JOURNAL

REFERENCE

AUTHORS

2 Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotseuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.

NEDO human cDNA sequencing project

TITLE

Unpublished

JOURNAL

REFERENCE

AUTHORS

Isogai,T. and Otsuki,T.

TITLE

Direct Submission

JOURNAL

Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

source

1. 2361
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3"

mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 53.1 Length: 2361
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x AK074790 (1-2361)

Qy 1 SerLeuGlnThrSerTyValPheLeu 9

Db 1476 TCCCTGCAAACTCTTATGCTTCCTG 1502

RESULT 29

BC087011

LOCUS BC087011 2361 bp mRNA linear ROD 13-DEC-2004
DEFINITION Rattus norvegicus trophoblast glycoprotein, mRNA (cDNA clone
MGC:93332 IMAGE:7193411), complete cds.

ACCESSION

BC087011

VERSION BC087011.1 GI:56268819

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

1 Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,P.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE

JOURNAL

PUBMED

12477932

2 (bases 1 to 2361)

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (02-DEC-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590,
USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 186 Row: 0 Column: 24
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

source

```

1..2361
  /organism="Rattus norvegicus"
  /mol_type="mRNA"
  /db_xref="taxon:10116"
  /clone="MGC:93332 IMAGE:7193411"
  /tissue_type="Heart, rat (Brown Norway)"
  /clone_lib="NIH MGC_234"
  /lab_host="DH10B"
  /note="Vector: pExpress1"
1..2361
  /gene="Tpbg"
  /note="synonym: 574"
  /db_xref="GeneID:83684"
  /db_xref="RGD:621453"
364..1644
  /gene="Tpbg"
  /codon_start=1
  /product="Tpbg protein"
  /protein_id="AAH87011.1"
  /db_xref="GI:56268820"
  /db_xref="GeneID:83684"
  /db_xref="RGD:621453"
  /translation="MPGAGSRGSPAGDGLRLRLALVLLGWVSAPSSSIPSSSTS
  PAAPLACGAPPPAPADPCAAECSEAAATKVCNRNLEVPADLPVVRNLFLLGNQ
  MTVLPAAGAPAPPLADLANLSCNHLKEVGAGAFELGLRLDLSHNPLTNLSAP
  TPAGNSVSTPSELLELILNHIIVPEQQRNGSFEGMVAFEGWVAALRGGLALRG
  HLELASHFLLPRDLQLPSLKHLDNRNLSVLSLYASFRNTHLESGLHLEADL
  KVLHNSTLARQGLAHVRVFLDNNPWCDCYMDVSWLKETEVPVDPKARLTCAPEK
  MNRGLLDLTSDLDLDCATLPQSQTSTVFLGLVLAIGALFLLVLYLNKGIKKKWH
  NIRDACRDHMGYHYRYINADPLRTNLSNSDV"
```

ORIGIN

Alignment Scores:
 Pred. No.: 53.1 Length: 2361
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-7 (1-9) x BC087011 (1-2361)

Qy 1 SerLeuGlnThrSerTyValPheLeu 9

Db 1432 TCCCTGCAGACTTCCTATGCTCTCTTA 1458

RESULT 30

BC037161 2379 bp mRNA linear PRI 29-JUN-2004
 LOCUS Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317
 IMAGE:4138906), complete cds.

ACCESSION

BC037161

VERSION BC037161.2 GI:33872201

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo

REFERENCE

1 (bases 1 to 2379)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carrinci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932

JOURNAL

PUBMED

2 (bases 1 to 2379)

Strausberg, R.

Direct Submission

TITLE

JOURNAL

Submitted (03-SEP-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USANIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:22713382.
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nigri.nih.gov
 Akhter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 26 Row: m Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5729717.

Location/Qualifiers

1..2379

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:15317 IMAGE:4138906"

/tissue_type="Muscle, rhabdomyosarcoma"

/clone_lib="NIH MGC_17"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1..2379

/gene="TPBG"

/note="Synonyms: M6P1, 574-AG, 574"

/db_xref="GeneID:7162"

/db_xref="MIM:190920"

427..1689

/gene="TPBG"

/codon_start=1

/product="574 oncofetal trophoblast glycoprotein"

/protein_id="AAH37161.1"

/db_xref="GI:22713383"

/db_xref="GeneID:7162"

/db_xref="MIM:190920"

/translation="MPGCSRGPAAGDGLRLRLALVLLGWVSSSPSSASSPSSSS"

APFLASAVSQPPLDQCPALCECSEAAATVKVNRNLTEVTDLPAYVNRNLFITGNQ
LAVLPAGAFARPPPLAEALANLSSGSLDEVRAGAEPHLPRLDLSNPLADLSPF
AFSGNSAVSAPSPVELLNLHIVPPEDRONRSPFGMVAAALAGLQGLRLLELA
SHNPLFPRDVLQAPLSRLDLNNSLSVLTYSFNLTHLSLHEDNALKVLHNG
TLAEQLGPHIFRVLDDNNPWCDCHMADMTWLKETEYVQGDRLTCAYPEKGRNVL
LELNSADLDCDILPSPSLQTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMHNRDAC
RDHMEGYHYRYEINADPRLTNLSSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 53.4 Length: 2379
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x BC037161 (1-2379)

QY 1 SerLeuGlnThrSerTyValPheTeu 9

Db 1477 TCCCTGCAACCTCTTATGTCCTTCCTG 1503

RESULT 31

BC058198

LOCUS

DEFINITION BC058198 2423 bp mRNA linear ROD 21-OCT-2003
IMAGE:5353871, complete cds.

ACCESSION

BC058198

VERSION

BC058198.1

KEYWORDS

MGC.

SOURCE

Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

Klausner, R.D., Collins, P.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,
Hopkins, R.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheet, T.B., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loguella, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McGowan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE

JOURNAL

PUBMED

12477932

REFERENCE

STRAUSBERG, R.

TITLE

JOURNAL

SUBMITTED (15-SEP-2003)

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),
Gaithersburg, Maryland.

Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastriani, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 123 Row: p Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

source

1. 2423

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:68145 IMAGE:5353871"

/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating

ductal carcinoma. 5 month old virgin mouse."

/clone_lib="NCI CGAP_Mam6"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 2423

/genes="Tpbp"

/notes="synonym: 5T4"

/db_xref="GeneID:21983"

/db_xref="MGI:1341264"

402..1682

/genes="Tpbp"

/codon_start=1

/product="Tpbp protein"

/protein_id="AAH58198.1"

/db_xref="GI:34849574"

/db_xref="GeneID:21983"

/db_xref="MGI:1341264"

/translation="MFCNGSRGSGAGRLRLARLALVLGVWSASAPSSVPSSTTS
PFAFLASGAQPPAPACCECAATVCKVNRNLLEVPADLPYVNRNLFITGNQ
MTVLPAGAPARQPPALDLALNLSGNHLKEVCAGAPHLPLGLRLDLSNPLTNLSAF
APAGSNASVAPSPLLELILNHLVPPEDORONGSEFGWAPEGWVAALSLGLRGL
TCLLASNHFLPLPDLAQLPSRLYLDLRNLSIVLUTYASFNRLTHLSLHEDNAL
KVLHNSTLAEWQGLAHVKVPLDNNPWVDCYMDMVAWLKETEYVVPDKRLTCAPPEK
MNRNGLDLNSDLDCDAVLPSLQTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NTRDACRDHMEGYHYRYEINADPRLTNLSSNSDV"

misc_feature

642..1262

/genes="Tpbp"

/notes="COG4886; Region: COG4886, Leucine-rich repeat (LRR)

protein [function unknown]"

/db_xref="CDD:COG4886"

1299..1415

/genes="Tpbp"

/notes="LRRCT; Region: Leucine rich repeat C-terminal

domain"

/db_xref="CDD:smart00082"

ORIGIN

Alignment Scores:

Pred. No.: 54.2 Length: 2423
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-7 (1-9) x BC058198 (1-2423)

```
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1470 TCCCTGCAGACTTCTCTATGCTTCTCTA 1496

RESULT 32
AX961912 2557 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 123 from Patent WO03104277.
ACCESSION AX961912
VERSION AX961912.1 GI:40881322
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Sugahara, T., Matsuura, A., Honda, G., Muramatsu, S. and Ishizawa, K.
Stat6 activation gene
Patent: WO 03104277-A 123 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1..2557
/db_xref="taxon:10090"
/molecule="Mus musculus"
/note="unassigned DNA"
556..1836
/note="unassigned protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881322"
/translation="MPGAGSGPSAGDGRRLRLALVLLGWVSASAPSSVPSSTTS
PADFLASGAQPPAERCPAACESEAAATVKCVNRNLLVPAADLPFYVRLFLTGNG
MTVLPAGAPAPPLADLEALNLSGNHLKEVCAGAFELPGLRLDLNHNPLNLISAF
VFAGNSASVSPSPLEELILNHIYPEDQKQSGFEGMVAFAALRSLALRGL
TRLKSLNHLFLPRDLIAQLPSRLYDLRNNLSVLTYSFRNLTHLSLHLEDNAL
KVLHNSLTAEWGLAHVKVFLDNNPMVCDYMDVMVWLKETEVPDKARLTCAPEK
MRNGLDLNSDLDLDCDAVLPSQLTSYVFLGIVLALIGALFLLVLYLNKGIKKMWH
NIRDACRDHMEGYHYRYEINADPRNLNLSNSDV"
ORIGIN
Alignment Scores:
Pred. No.: 56.8 Length: 2557
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-7 (1-9) x AX961912 (1-2557)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1624 TCCCTGCAGACTTCTCTATGCTTCTCTA 1650

RESULT 34
AX168308 2714 bp mRNA linear PRI 18-JUN-2005
LOCUS
DEFINITION Macaca fascicularis testis cDNA clone: QcSA-11109, similar to human
trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.
ACCESSION AB168308
VERSION AB168308.1 GI:67967899
KEYWORDS
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE
1 International consortium for macaque cDNA sequencing and analysis.
AUTHORS DNA sequences of macaque genes expressed in brain or testis and its
TITLE evolutionary implications
JOURNAL
REFERENCE Unpublished
AUTHORS
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
```

Biology, National Institute of Genetics of Japan, Mishima, Japan.
Clone distribution: clone distribution information can be found at:
<http://www.nih.go.jp/yoken/genebank/>

Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CAGCTGTGTG)
R. Site2: DraIII (CAGCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by oligo-capping method. Libraries were made from:

- Qccs: cerebellum cortex
- QnpA: parietal lobe
- QnpB: temporal lobe right
- Qf1A: frontal lobe left
- QmoA: medulla oblongata
- QbsA: brain stem
- QorA: occipital lobe right
- QtsA: testis

Custom primers were used for 5' and 3'-end sequencing. The
full-insert sequencing was done by primer-walking method using ABI
DNA sequencer.

Location/Qualifiers
1..2714
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="QtsA-11109"
/sex="male"
/clone_lib="macaque cDNA library QtsA"
/dev_stage="adult"
764..2026

/note="unnamed protein product; Homo sapiens trophoblast
glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3"
/codon_start=1
/protein_id="BAE00432.1"
/db_xref="GI:67967900"

/translation="MPGCSGPGAGDGLRLRLARLALVLLGWSSSTSSASSSSSS
APFLASASAPPLPDQPCALCESEARTVKVNRNLTEVPTDLPVLRNPLFTGNQ
LAVLPAGAFARPPPLAEALNLGSRSLDEVRAGAFELHPSLRQLDLSHNPGLADLSPF
AFSGSNASVSPPLVELLNHIVPDEQRNRSFEGMVVAALLAGRALQGLRLLELA
SNHFLYLPDVLQPLSLRLDLSNNLSVLTYSFRNLTHLSLHLEDNALKVLHNG
TLAEIQLGPHVRVLDNNPWVCDCHMADMTWLKQTVVQGDRLTCAPEKWRNRL
LELNSADLDCDPIPLPSPQTSYVFLGIVLALIGALFLVLVLRNKGKIKKWMHNRDLC
RDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
Pred. No.: 59.7 Length: 2714
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x ABL69308 (1-2714)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1814 TCCCTGCAAACTCTTATGCTCTCTG 1840

RESULT 35
HSA012159
LOCUS HSA012159 5551 bp DNA linear PRI 15-APR-2005
DEFINITION Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AJ012159
VERSION AJ012159.1 GI:3805946
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1
King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues

JOURNAL

Biochim. Biophys. Acta 1445 (3), 257-270 (1999)

PUBMED

10366710

REFERENCE

2 (bases 1 to 5551)

AUTHORS

Myers, K.A.

TITLE

Direct Submission

JOURNAL

Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK

FEATURES

Location/Qualifiers

1..5551

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

2698..2703

/bound_moiety="Sp1"

2704..2709

/bound_moiety="Sp1"

2716..5400

/gene="5T4"

2716..2800

/gene="5T4"

/evidence="experimental"

2801..3092

/gene="5T4"

/evidence="experimental"

3093..5400

/gene="5T4"

/evidence="experimental"

3431..54693

/gene="5T4"

/codon_start=1

/product="5T4 oncofetal trophoblast glycoprotein"

/protein_id="CAA09930.1"

/db_xref="GI:3805947"

/db_xref="GOA:Q13641"

/db_xref="InterPro:IPR000372"

/db_xref="InterPro:IPR000483"

/db_xref="InterPro:IPR001611"

/db_xref="InterPro:IPR003591"

/db_xref="UniProt/TREMBL:Q13641"

/translation="MPGCSGPGAGDGLRLRLARLALVLLGWSSSTSSASSSSSS
APFLASASAPPLPDQPCALCESEARTVKVNRNLTEVPTDLPVLRNPLFTGNQ
LAVLPAGAFARPPPLAEALNLGSRSLDEVRAGAFELHPSLRQLDLSHNPGLADLSPF
AFSGSNASVSPPLVELLNHIVPDEQRNRSFEGMVVAALLAGRALQGLRLLELA
SNHFLYLPDVLQPLSLRLDLSNNLSVLTYSFRNLTHLSLHLEDNALKVLHNG
TLAEIQLGPHVRVLDNNPWVCDCHMADMTWLKQTVVQGDRLTCAPEKWRNRL
LELNSADLDCDPIPLPSPQTSYVFLGIVLALIGALFLVLVLRNKGKIKKWMHNRDLC
RDHMEGYHYRYEINADPRLTNLSNSDV"

3431..3516

/gene="5T4"

3517..54690

/gene="5T4"

/product="5T4 oncofetal trophoblast glycoprotein"

5331..5336

/gene="5T4"

5380..5385

/gene="5T4"

ORIGIN

Alignment Scores:

Pred. No.: 109 Length: 5551

Score: 43.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0


```

Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x HSA012159 (1-5551)

Qy 1 SerLeuGlnThrSerTyValpHeLeu 9
   |||||
Db 4481 TCCCTGCAGACTCTCTATGCTCTCCTG 4507

RESULT 36
MMU012160 7942 bp DNA linear ROD 15-APR-2005
LOCUS Mus musculus 5T4 oncofetal trophoblast glycoprotein gene.
DEFINITION
ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10366710
REFERENCE 2 (bases 1 to 7942)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK

FEATURES
source
1..7942
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /strain="129/Sv"
   /db_xref="taxon:10090"
   /clone_lib="Lambda Dash"
   3108..3113
   /bound_molty="Spl"
   3114..3119
   /bound_molty="Spl"
   3124..5779
   /gene="5T4"
   3124..3151
   /gene="5T4"
   3152..3450
   /gene="5T4"
   3451..5779
   /gene="5T4"
   3779..5059
   /gene="5T4"
   /codon_start=1
   /product="5T4 oncofetal trophoblast glycoprotein"
   /protein_id="CAA09931.1"
   /db_xref="GI:3805949"
   /db_xref="GOA:Q920L0"
   /db_xref="InterPro:IPR000372"
   /db_xref="InterPro:IPR000483"
   /db_xref="InterPro:IPR01611"
   /db_xref="InterPro:IPR003591"
   /db_xref="MGI:1341264"
   /translation="MPGAGSGPAGDGRLLRLRLALVLLGWVSASAPSSVPSSTTS
PADFLAGSAGPPAPRCFAACESEARTKVCNRLLEVPADLPFYRNILFTGNQ
MTVLPAFAFAPQPLADLEALNSGNHLKVCAGAFHLPGLRLDLSHNLPLTNSAF
TFAGSNASVAPSPLLEILNHIVPPDQRQNGSPGVAPEGVAAALRSGLAGRL
VRLASNHPFLPLRLLAQLPSRLYLDLRNNSLVSLTYASFRLNLTLESLELDNAL
KYLHNSILAEOGLAHVKYFLDNPWCDCYMDAMVWLKETEVPDKARLTCAPEK
MNRNGLDLNSSLDDCDVLFQSLQTSYFLVGLVIALIGALFLVLYLNKGIKKMKH

```

```

sig_peptide 3779..3865
mat_peptide 3866..5056
polyA_signal 5713..5718
polyA_signal 5759..5764

NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"
/gene="5T4"
3866..5056
/gene="5T4"
/product="5T4 oncofetal trophoblast glycoprotein"
5713..5718
/gene="5T4"
5759..5764
/gene="5T4"

ORIGIN
Alignment Scores: 148 Length: 7942
Pred. No.: 43.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 9
US-10-774-176-7 (1-9) x MMU012160 (1-7942)

Qy 1 SerLeuGlnThrSerTyValpHeLeu 9
   |||||
Db 4847 TCCCTGCAGACTCTCTATGCTCTCCTA 4873

RESULT 37
HSJ492P14 121909 bp DNA linear PRI 18-MAY-2005
LOCUS Human DNA sequence from clone RP3-492P14 on chromosome 6q13-15
DEFINITION Contains a single stranded DNA binding protein pseudogene, the TPBG
gene for trophoblast glycoprotein (5T4-AG) and a CpG island,
complete sequence.
ACCESSION ALI21977
VERSION ALI21977.11 GI:11863678
KEYWORDS HTG; CpG island; TPBG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
          1 (bases 1 to 121909)
          Garner, P.
          Direct Submission
          Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
          Clone requests: clonerequest@sanger.ac.uk
          On Dec 15, 2000 this sequence version replaced gi:11558491.
          The following abbreviations are used to associate primary accession
          numbers given in the feature table with their source databases:
          Em., EMBL; SWI, SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
          on the WORMPEP database can be found at
          http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
          was generated from part of bacterial clone contigs of human
          chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
          Group. Further information can be found at
          http://www.sanger.ac.uk/HGP/Chr6
          RP3-492P14 is from the library RPCI-3 constructed by the group of
          Pieter de Jong. For further details see
          http://www.chori.org/bacpac/home.htm
          VECTOR: pCYPAC2
          ----- Genome Center
          Center: Wellcome Trust Sanger Institute
          Center code: SC
          Web site: http://www.sanger.ac.uk
          Contact: vegas@sanger.ac.uk
          -----
          This sequence was finished as follows unless otherwise noted: all
          regions were either double-stranded or sequenced with an alternate
          chemistry or covered by high quality data (i.e., phred quality >=
          30), an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by at least
          one subclone, and the assembly was confirmed by restriction digest,

```

except on the rare occasion of the clone being a YAC.

FEATURES

source

```
1. 121909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="q13-15"
/clone="RP3-492P14"
/clone_lib="RPCI-3"
```

misc_feature

```
100
/note="Clone right_end: RP1-93K22"
complement(10004..10982)
/locus_tag="RP3-492P14.2-001"
```

gene

```
/pseudoc
complement(10004..10982)
/locus_tag="RP3-492P14.2-001"
```

CDS

```
/note="match: proteins: P81877 Q99LX9 Q9BWM6 Q9CYZ8 Q9D6L4
Q9P038 Q9Y4T7"
```

misc_feature

```
/pseudoc
/codon_start=1
```

gene

```
86539
/note="Clone left_end: RP1-90G1"
109639..116836
/gene="TPBG"
```

mRNA

```
/locus_tag="RP3-492P14.1-001"
join(109639..109916,110631..116836)
/gene="TPBG"
```

```
/locus_tag="RP3-492P14.1-001"
/product="trophoblast glycoprotein"
```

```
/note="match: ESTs: AA149121 AA152323 AA565852 AA643734
```

```
AL544610 AW471072 AW662538 BE260089 BF306457 BF306926
```

```
BF314984 BI196133 BI562387 BM069633 BM670613
```

```
match: CDNAS: AJ420536.1 Z29083.1"
```

CDS

```
110970..112232
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
/standard_name="OTHUMP0000016786"
```

```
/note="match: proteins: Q13641 Q9QXD9 Q9Z0L0"
```

```
/codon_start=1
```

```
/product="trophoblast glycoprotein"
```

```
/protein_id="CAI21546.1"
```

```
/db_xref="GI:56203539"
```

```
/db_xref="Genew:12004"
```

```
/db_xref="GOA:Q13641"
```

```
/db_xref="InterPro:IPR000372"
```

```
/db_xref="InterPro:IPR000483"
```

```
/db_xref="InterPro:IPR001611"
```

```
/db_xref="InterPro:IPR003591"
```

```
/db_xref="UniProt/TREMBL:Q13641"
```

```
/translation="MPGCSRGPAAGDGRRLRLALVLLGWSSSPTSSASSPSS
```

```
APFLASVSAQPLPDQCPALCECEAAATVKVNRNLTEVPTDLPAYVRLPLTGNQ
```

```
LAVLPAGAFAPRLAEALNLSGSLDEVRAGAEHLPSLRQLDLSHNPDLSPF
```

```
AFSGSNASVSPSLVELILNHIVPEDERQNSPEGMVAALLAGRLRLRLA
```

```
SNHFLYLRPLDQLPSLRHLDSNLSVLTYSFRNLTHLSLHLEDAKLVKNG
```

```
TLAAQGFPHRIVFDNPNWCDCHMADWTLKTEVVQKDRLTCAYPEKRRVL
```

```
LELNADLDCDPLPSLQTSYVFLGIVLALIGAFLLVLYLNKRGKIKKMMNIRDAC
```

```
RDMGEGYHYRYEINADPRLTNLSSNDV"
```

```
116817..116822
```

polya_signal

```
/gene="TPBG"
```

```
/locus_tag="RP3-492P14.1-001"
```

```
116836
```

polya_site

```
/gene="TPBG"
```

```
/locus_tag="RP3-492P14.1-001"
```

misc_feature

```
121909
/note="Clone_right_end: RP3-492P14"
```

ORIGIN

Alignment Scores: 1.49e+03 Length: 121909

Pred. No.: 43.00 Matches: 9

Score: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x HSJ492P14 (1-121909)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9

DB 112020 TCCTGCAAACTCTTATGTCCTTCCTG 112046
|||||

RESULT 38

AC158516/C

LOCUS

DEFINITION

AC158516 AC117768

AC158516.2 GI:63025421

AC158516.2 GI:63025421

HTG.

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

Adams,S., Cotton,M. and Haglund,K.

The sequence of Mus musculus BAC clone RP24-511A23

Unpublished (2001)

2 (bases 1 to 167046)

Wilson,R.K.

Direct Submission

Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park

Parway, St. Louis, MO 63108, USA

3 (bases 1 to 167046)

Wilson,R.K.

Direct Submission

Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park

Parway, St. Louis, MO 63108, USA

4 (bases 1 to 167046)

Wilson,R.K.

Direct Submission

Submitted (21-JUN-2005) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On May 4, 2005 this sequence version replaced gi:61656412.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@wustl.edu

----- Summary Statistics

Center project name: M_BB0511A23

Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e. phred quality
>30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone, fosmid clone or direct clone walk sequence.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to obtain the consensus sequence. The
assembly was confirmed by restriction digest.

This finishing standard has slightly changed from the previous
Human standard. Specifically, standards for regions of low sequence
complexity (such as dinucleotide repeats and small unit tandem
repeats) have been relaxed. These regions are very prevalent in the
mouse genome, and the return on extended finishing efforts is
minimal.

If a sequence meets the criteria of the above statement, it needs
no comments or tags. If the criteria are not met, such as ambiguous
bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

```

FEATURES             source
    Location/Qualifiers
        1..167046
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="9"
            /clone="RP24-511A23"
            /clone_lib="RPCI-24"
        misc_feature
            16685..16712
                /note="Sequence derived from PCR product of genomic DNA"
        unsure
            31565..31779
                /note="Unresolved simple sequence repeat."
        unsure
            46721..46808
                /note="Unresolved simple sequence repeat."
        unsure
            142336..142347
                /note="Sequence derived from one plasmid subclone."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      1.94e+03      Length:      167046
Score:          43.00         Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      100.0%         Indels:      0
DB:              9             Gaps:      0

```

US-10-774-176-7 (1-9) x AC158516 (1-167046)

Qy 1 SerLeuGlnThrSerTyrValpheLeu 9

Db 109770 TCCTGCGAGACTTCTATGCTTCCTA 109744

```

RESULT 39
AC128294/c
LOCUS      AC128294      210237 bp      DNA      linear      HTG 19-NOV-2002
DEFINITION Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.
ACCESSION AC128294
VERSION   AC128294.3 GI:25083347
KEYWORDS  HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 210237)
Muzny D.Marie, Metzker M.Lee, Abramson S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Ballwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed P.,
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chaver D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.H., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Evans K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.N., Gabisi A., Ganta R., Garcia A., Garner I., Garza M.,

```

Geбреgeorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W., Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K., Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J., Hernandez R., Hines S., Hladun S.L., Hodgson A., Hognes M., Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A., Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A., Karpasch S., Kelly S., Kelly S., Khan Z., King L., Kovar C., Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J., Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J., Lorenshewa L., Louisaeged H., Lozado R.J., Lu X., Ma J., Mangum A., Mangum B., Mapua P., Martin K., Martin R., Malloy K., Manguam A., Mawhiney S., McLeod M.P., McNeill T.Z., Meenen E., Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S., Morgan M., Morris K., Morris S., Munidase M., Murphy M., Nair L., Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Nair L., Nwakoelameh O., Okwuonu G., Olarnpunaagoon A., Pal S., Parks K., Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C., Plopper F., Poindexter A., Popovic D., Primus E., Pu L., Pu L., Puazo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R., Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs P., Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J., Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajs D., Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J., Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C., Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Umanik K., Valas R., Vera V., Villalana D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren R., Warren R., Wei X., White F., Williams G., Willson R., Wleczyk R., Wooden H., Worley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu P., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL

COMMENT

Unpublished
2 (bases 1 to 210237)

Worley K.C.

Direct Submission

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 210237)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265004.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: CH230-176H20
Center clone name: CH230-176H20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 201781 bases at least Q40
Consensus quality: 203921 bases at least Q30

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GOPI

Center clone name: CH230-87110

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 228642 bases at least Q40

Consensus quality: 232269 bases at least Q30

Consensus quality: 234041 bases at least Q20

Estimated insert size: 231522; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

consists of 4 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

* 1 234710: contig of 234710 bp in length

* 234711 234810: gap of unknown length

* 234811 235924: contig of 1114 bp in length

* 235925 236024: gap of unknown length

* 236025 237314: contig of 1290 bp in length

* 237315 237414: gap of unknown length

* 237415 239076: contig of 1662 bp in length.

Location/Qualifiers

1. .239076

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-87110"

234711..234810

/estimated_length=unknown

235925..236024

/estimated_length=unknown

237315..237414

/estimated_length=unknown

ORIGIN

Alignment Scores:

Pred. No.: 2.63e+03 Length: 239076

Score: 43.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 14 Gaps: 0

US-10-774-176-7 (1-9) x AC106962 (1-239076)

Oy 1 SerLeuGlnThrSerTyrValPheLeu 9

|||||

Db 15661 TCCTGCAGACTTCCTATGCTCTCTTA 15635

RESULT 41

BC099984

LOCUS

DEFINITION

Danio rerio trophoblast glycoprotein-like, mRNA (CDNA clone

MGC:109764 IMAGE:7226481), complete cds.

ACCESSION

BC099984

VERSION

BC099984.1 GI:71534092

KEYWORDS

MGC.

SOURCE

Danio rerio (zebrafish)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 1713)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richard, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

CONSR TM

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

CONSR TM

TITLE

JOURNAL

REMARK

COMMENT

NIH MGC Project
Direct Submission
Submitted (29-JUL-2005) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Will Talbot
CDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 210 Row: f Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 35902948.

FEATURES

Source

Location/Qualifiers

1..1713

/organism="Danio rerio"

/mol_type="mRNA"

/db_xref="taxon:7955"

/clone="MGC:109764 IMAGE:7226481"

/tissue_type="Embryo, whole, 2-8 hr, mixed wild type, pool

of 2500 embryos"

/clone_lib="NIH_ZGC_15"

/lab_host="DH10B"

/note="Vector: pME18S-FL3"

1..1713

/gene="tpbgl"

/note="synonyms: ab:cb613, YF-296"

/db_xref="GeneID:373095"

/db_xref="ZFIN:ZDB-GENE-030827-4"

57..1175

/gene="tpbgl"

/codon_start=1

/product="trophoblast glycoprotein-like"

/protein_id="AAH99984.1"

/db_xref="GI:71534093"

/db_xref="GeneID:373095"

/db_xref="ZFIN:ZDB-GENE-030827-4"

/translation="MPARGCAVLGLLCAAAVAGALVPTGCECEAAITVKCVSK

DLRDIPSGIPGYTRNLFTGNHISQIGPESFQLENVTNLSLNRISEVKSTFSSL

CDS

RSLSRLDLSNNLAVIHPAFTVQSRMLRELNSRALYNHSSVMDLATSLRWSSLSDL
LVLDLSSNGLVLPSCIFCHVGLRELQIAGNSIVSIHNGTFTGLDLRLQELDLTHNAL
PRLREALXELQLSARLHLADNPFTCTCDIEPPAAMNGSRGVQVVDIEGLTCAPPV
ALHNTSLTVGDLGCKHAGSDNLALQTSIVFVLGVLGFLVGLMFLFVLYLNRKDIK
KKIYDMRDACREVWEGYHYRYEIDSDPRLSQVSSTADV"

ORIGIN

Alignment Scores:
Pred. No.: 168 Length: 1713
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-7 (1-9) x BC099984 (1-1713)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 963 GCGCTACAGACCTCTTACGTTCTCTG 989

RESULT 42

AB097825 1765 bp mRNA linear VRT 11-JUN-2003
LOCUS
DEFINITION
Danio rerio YF-296 mRNA for hypothetical protein, complete cds.
ACCESSION
AB097825
VERSION
AB097825.1 GI:31580869
KEYWORDS
Danio rerio (zebrafish)
SOURCE
ORGANISM

Yoda,H., Momoi,A., Begueria,C., Meyer,D., Driever,W., Kondoh,H. and

Furutani-Seiki,M.
An expression pattern screen for genes involved in induction of the
posterior nervous system of zebrafish
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 1765)
Yoda,H., Kondoh,H. and Furutani-Seiki,M.
Direct Submission
Submitted (10-DEC-2002) Makoto Furutani-Seiki, ERATO Kondoh
Differentiation Signaling Project, JST; Kawara-machi 14, Yoshida,
Sakyo-Ku, Kyoto 606-8305, Japan
(E-mail: furutani-seiki@msi.biglobe.ne.jp, Tel:81-75-771-9362,
Fax:81-75-771-8281)
Location/Qualifiers
1..1765
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"

1..1765
/gene="YF-296"
169..1287
/gene="YF-296"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC77539.1"
/db_xref="GI:31580870"

/translation="MFARGRCVAVLGLLCAAAVSAGALVCTGCECEAAITVKCVSK
RLRDTPSGTGYTRNLFTGNHISQIPESFQGLNVTNLSNNRISVKGHTFSSL
DLSRLDLSNNOLAVIHPAFTVQSRMLRELNSRALYNHSSVMDLATSLRWSSLSDL
LVLDLXSLGLVPLPSGIFCHVGLRELQIAGNSIVSIHNGTFTGLDLRLQELDLTHNAL
PRLREALXELQLSARLHLADNPFTCTCDIEPPAAMNGSRGVQVVDIEGLTCAPPV
ALHNTSLTVGDLGCKHAGSDNLALQTSIVFVLGVLGFLVGLMFLFVLYLNRKDIK
KKIYDMRDACREVWEGYHYRYEIDSDPRLSQVSSTADV"

Alignment Scores:
Pred. No.: 173 Length: 1765
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1

Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-7 (1-9) x AB097825 (1-1765)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1075 GCGCTACAGACCTCTTACGTTCTCTG 1101

RESULT 43

AL353584 2/c
WPCOMMENT
Sequence split into 6 fragments LOCUS AL353584 Accession AL353584
Fragment Name Begin End
AL353584_0 1 110000
AL353584_1 100001 210000
AL353584_2 200001 310000
AL353584_3 300001 410000
AL353584_4 400001 510000
AL353584_5 500001 536214
Continuation (3 of 6) of AL353584 from base 200001 (AL353584 Homo sapiens chromosome X)

Alignment Scores:
Pred. No.: 5.67e+03 Length: 110000
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-7 (1-9) x AL353584_2 (1-110000)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 82851 TCTCTGCAACTAGCTATCTATTTCTG 82825

RESULT 44

AL445194/131711 bp DNA linear HTG 10-JUL-2001
LOCUS
DEFINITION
Homo sapiens chromosome X clone RP11-466G5, 17 unordered pieces.
ACCESSION
AL445194
VERSION
AL445194.2 GI:12539754
KEYWORDS
HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1
Mclay,K.
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jan 26, 2001 this sequence version replaced gi:10716446.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA466G5
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator ET-amersham; 8% of reads Chemistry:
Dye-terminator Big Dye; 91% of reads
Consensus quality: 123858 bases at least Q40
Consensus quality: 127031 bases at least Q30
Consensus quality: 128895 bases at least Q20
Insert size: 130111; sum-of-contigs
Insert size: 166918; 7.0% error; agarose-fp

Quality coverage: 3.46x in Q20 bases; sum-of-contigs Quality coverage: 2.78x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 12531: contig of 12531 bp in length
12532: gap of 100 bp
12632: contig of 4825 bp in length
17457: gap of 100 bp
17557: contig of 2158 bp in length
19714: gap of 100 bp
19715: contig of 7057 bp in length
19815: gap of 100 bp
26872: gap of 100 bp
26971: contig of 4664 bp in length
31635: gap of 100 bp
31735: gap of 100 bp
31736: contig of 5584 bp in length
37320: gap of 100 bp
37420: contig of 15824 bp in length
53243: gap of 100 bp
53244: gap of 5409 bp in length
58752: gap of 100 bp
58853: contig of 3713 bp in length
62565: gap of 100 bp
62666: contig of 4605 bp in length
67270: gap of 100 bp
67271: contig of 8081 bp in length
75451: gap of 100 bp
75551: contig of 15018 bp in length
90570: gap of 100 bp
90670: contig of 5694 bp in length
96363: gap of 100 bp
96464: contig of 11735 bp in length
108199: gap of 100 bp
108299: contig of 11268 bp in length
119567: gap of 100 bp
119667: contig of 2643 bp in length
122310: gap of 100 bp
122410: contig of 9302 bp in length.

FEATURES
source

1..131711
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-466G5"
/clone_lib="RPC1-11.2"
1..12531
/note="assembly_fragment:00689
clone_end:17
vector_side:left"
12632..17456
/note="assembly_fragment:00069
fragment_chain:1"
17557..19714
/note="assembly_fragment:00751
fragment_chain:1"
19815..26871
/note="assembly_fragment:00859
fragment_chain:1"
26972..31635
/note="assembly_fragment:00573
fragment_chain:1"
31736..37319
/note="assembly_fragment:01277
fragment_chain:2"
37420..53243
/note="assembly_fragment:00700

misc_feature
fragment chain:2"
53344..58752
/note="assembly_fragment:00257
fragment chain:2"
58853..62565
/note="assembly_fragment:00728
fragment chain:2"
62666..67270
/note="assembly_fragment:01136
fragment chain:3"
67371..75451
/note="assembly_fragment:00624
fragment chain:3"
75552..90569
/note="assembly_fragment:00196.0"
90670..96363
/note="assembly_fragment:00196.1"
96464..108198
/note="assembly_fragment:00503"
108299..119566
/note="assembly_fragment:00528"
119667..122309
/note="assembly_fragment:00442
fragment chain:4"
122410..131711
/note="assembly_fragment:01264
fragment chain:4
clone_end:SP6
vector_side:right"

ORIGIN

Alignment Scores:
Pred. No.: 6.6e+03 Length: 131711
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-7 (1-9) x AL445194 (1-131711)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
|||||
Db 103495 TCTCTGCAACTAGCTACTATTCTG 103469

RESULT 45

AL139114
LOCUS Human DNA sequence from clone Rp11-328C23 on chromosome 9p22.3-24.1, complete sequence.
DEFINITION
ACCESSION AL139114
VERSION AL139114.12 GI:10045316
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 132139)
AUTHORS Sehra, H.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT
On Sep 9, 2000 this sequence version replaced gi:9943980.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TRAMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at

```

http://www.sanger.ac.uk/HGP/Chr9
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
RP11-328C23 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES
Source
1..132139
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="9"
    /map="p22.3-24.1"
    /clone="RP11-328C23"
    /clone_lib="RPCI-11.2"
1
    /note="Clone_left_end: RP11-328C23"
44274
    /note="Clone_right_end: RP11-364M22"
132040
    /note="Clone_left_end: RP11-3L8"

misc_feature
1
    /note="Clone_left_end: RP11-328C23"

misc_feature
44274

misc_feature
132040

ORIGIN
Alignment Scores:
Pred. No.: 6.62e+03 Length: 132139
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x AL139114 (1-132139)
Oy 1 SerLeuGlnThrSerTyrValPheLeu 9
:::|||||
Db 97496 ACTTGGCAACCGTATGATTTCTC 97522

RESULT 46
AC119204
LOCUS AC119204 178603 bp DNA linear ROD 02-JUN-2005
DEFINITION Mus musculus chromosome 1, clone RP24-188H2, complete sequence.
ACCESSION AC119204
VERSION AC119204.7 GI:66865113
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 178603)
Biren,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 1, clone RP24-188H2
Unpublished
2 (bases 1 to 178603)
Biren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kelle,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Mathews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retka,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,
DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kelle,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Mathews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Nicol,R., Norbu,C.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retka,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 178603)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,
DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kelle,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Mathews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Nicol,R., Norbu,C.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retka,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-MAR-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 178603)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepel,Y., Collymore,A., Cooke,A., Cooke,P., Corum,B.,
DeArrellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kelle,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Mathews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retka,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (02-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
On Jun 2, 2005 this sequence version replaced gi:61696444.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

```


Center: Broad Institute of MIT and Harvard
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@broad.mit.edu
 ----- Project Information
 Center project name: I25193
 Center clone name: 188_H_2

FEATURES

```

source
1..178603
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /chromosome="1"
  /map="1"
  /clone="RP24-188H2"
  /clone_lib="RPCI-24 Male Mouse BAC"
  381..422
    /rpt_family="L2"
    complement(964..1040)
    /rpt_family="B4"
    repeat_region
    complement(1091..1251)
    /rpt_family="B3A"
    repeat_region
    complement(1252..1465)
    /rpt_family="B4"
    repeat_region
    complement(1601..1719)
    /rpt_family="BCL_MM"
    repeat_region
    complement(1631..1722)
    /rpt_family="ID_B1"
    repeat_region
    2503..2609
    /rpt_family="PB1D10"
    repeat_region
    2627..2703
    /rpt_family="GAA"
    repeat_region
    2704..2883
    /rpt_family="GAA"
    repeat_region
    2981..3232
    /rpt_family="B4"
    repeat_region
    3358..3502
    /rpt_family="B1P"
    repeat_region
    3514..3580
    /rpt_family="GAAA"
    repeat_region
    3590..3773
    /rpt_family="B3"
    complement(5033..5160)
    /rpt_family="ID_B1"
    repeat_region
    5191..5272
    /rpt_family="PB1D7"
    complement(6564..6672)
    /rpt_family="ID_B1"
    repeat_region
    6785..6928
    /rpt_family="B1_MM"
    repeat_region
    6929..6952
    /rpt_family="CAAC"
    repeat_region
    7243..7265
    /rpt_family="CAGTT"
    repeat_region
    complement(7321..7474)
    /rpt_family="B1_MM"
    repeat_region
    7520..7540
    /rpt_family="AT_rich"
    repeat_region
    7901..8064
    /rpt_family="RSINE1"
    complement(8083..8204)
    /rpt_family="RSINE1"
    repeat_region
    8522..8568
    /rpt_family="GA"
    repeat_region
    complement(8741..9124)
    /rpt_family="MTA"
    repeat_region
    complement(9397..9576)
    /rpt_family="ID_B1"
    repeat_region
    complement(9584..9684)
    /rpt_family="B1_MM"
    repeat_region
    complement(9711..10080)
    /rpt_family="MTD"

```

```

repeat_region complement(10088..10193)
/rpt_family="PB1D9"
repeat_region complement(10940..11043)
/rpt_family="L3"
repeat_region 11700..11882
/rpt_family="B2_Mm2"
repeat_region 12049..12093
/rpt_family="(TG)n"
repeat_region 13265..13317
/rpt_family="(TAAA)n"
repeat_region 13704..13925
/rpt_family="B3"
repeat_region 13959..14090
/rpt_family="B1_MM"
repeat_region 14623..14802
/rpt_family="B3"
repeat_region 15083..15166
/rpt_family="MERSA"
repeat_region complement(15323..15334)
/rpt_family="B1P"
repeat_region complement(15335..15411)
/rpt_family="ID_B1"
repeat_region complement(15412..15449)
/rpt_family="B1P"
repeat_region 15789..15823
/rpt_family="(CA)n"
repeat_region complement(15948..16077)
/rpt_family="B1_MM"
repeat_region complement(16201..16272)
/rpt_family="B1P"
repeat_region complement(16362..16466)
/rpt_family="B1_MM"

```

Alignment Scores:

```

Pred. No.: 8.54e+03 Length: 178603
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 9 Gaps: 0

```

US-10-774-176-7 (1-9) x AC119204 (1-178603)

```

QY 1 SerLeuGlnThrSerTyValPheIeu 9
|||||:::|||||
Db 169358 TCTCTACAGACCACTATGTTTTTA 169384

```

RESULT 47

```

AC116853
LOCUS AC116853 188514 bp DNA linear ROD 19-OCT-2004
DEFINITION Mus musculus chromosome 1, clone RP24-400M20, complete sequence.
ACCESSION AC116853
VERSION AC116853.12 GI:54291915
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

```

```

REFERENCE 1 (bases 1 to 188514)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus chromosome 1, clone RP24-400M20
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 188514)
AUTHORS

```

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faros,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (02-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 188514)

REFERENCE
AUTHORS

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (28-JUL-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 188514)

REFERENCE
AUTHORS

Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (19-OCT-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 19, 2004 this sequence version replaced gi:50726824.

COMMENT

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu

----- Project Information

Center project name: L25620
Center clone name: 400_M20

FEATURES

source

Location/Qualifiers

1..188514
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"
/clone="RP24-400M20"
/clone_lib="RPCI-24 Male Mouse BAC"
1..443
/rpt_family="Lx8"
complement(516..566)
/rpt_family="B2_Mm2"
complement(632..804)
/rpt_family="B2_Mm1"
complement(874..1020)
/rpt_family="B1_MM"
complement(1178..1410)
/rpt_family="B4A"
complement(1429..1622)
/rpt_family="HAL1"
1776..1824
/rpt_family="(TTCC)n"
complement(1825..1838)
/rpt_family="B4A"
complement(1839..1957)
/rpt_family="PBID9"
complement(1958..1998)
/rpt_family="B4A"
1999..2037
/rpt_family="(TG)n"
complement(2086..2655)
/rpt_family="Lx2"
2655..2810
/rpt_family="L1"
2811..2855
/rpt_family="(CAAA)n"
2856..2947
/rpt_family="L1"
2948..2983
/rpt_family="(CAAA)n"
2984..3211
complement(3277..3312)
/rpt_family="U2"
complement(3326..3443)
/rpt_family="B1_MM"
3466..3503
/rpt_family="AT_rich"
complement(3504..3690)
/rpt_family="B3"
complement(3691..3911)
/rpt_family="B3"
4054..4157
/rpt_family="ID_B1"
4277..4321
/rpt_family="(CA)n"
4437..4467
/rpt_family="AT_rich"
4691..4744
/rpt_family="(CAAA)n"
4951..4978
/rpt_family="(TTTG)n"
complement(4979..5112)
/rpt_family="B1_MM"
5128..5323
/rpt_family="B2_Mm2"

```
repeat_region 5417..5475
/rpt_family="(CA)n"
repeat_region 5498..5519
/rpt_family="AT-rich"
repeat_region complement(5520..5953)
/rpt_family="LIM4"
repeat_region complement(6282..6422)
/rpt_family="B1F"
repeat_region complement(6423..6468)
/rpt_family="B4"
repeat_region complement(6520..6837)
/rpt_family="MTE"
repeat_region complement(7338..7445)
/rpt_family="B1_MH"
repeat_region 7627..7651
/rpt_family="(CAAAA)n"
repeat_region complement(7885..7961)
/rpt_family="B4"
repeat_region complement(7962..8101)
/rpt_family="B1_MH"
repeat_region complement(8102..8126)
/rpt_family="B4"
repeat_region complement(8335..8571)
/rpt_family="L2"
repeat_region complement(8580..8708)
/rpt_family="L2"
repeat_region 10320..10436
/rpt_family="HAL1"
repeat_region 11170..11308
/rpt_family="WT2B"

Alignment Scores:
Pred. No.: 8.94e+03 Length: 188514
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.8% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-7 (1-9) x AC116853 (1-188514)
Qy 1 SerLeuGlnThrSerTyValpHeu 9
Db 23147 TCCTCAGACCACTATGTTTITA 23173

RESULT 48
AC019212 AC019212 213793 bp DNA linear PRI 09-MAY-2001
LOCUS Homo sapiens BAC clone RP11-462H12 from X, complete sequence.
DEFINITION AC019212
ACCESSION AC019212
VERSION AC019212.4 GI:9581967
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 213793)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
9847074
PUBMED
2 (bases 1 to 213793)
Grewal, N., Drone, K., Laplant, Y. and Le, T.
The sequence of Homo sapiens BAC clone RP11-462H12
Unpublished
3 (bases 1 to 213793)
Waterston, R.H.
Direct Submission
Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 213793)

repeat_region 1021..1146
/rpt_family="MIR"
repeat_region 1267..1366
/rpt_family="L2"

FEATURES
source
1..213793
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="RP11-462H12"
/clone_lib="RPCI-II"
137..322
/rpt_family="MIR"
1021..1146
/rpt_family="L2"
1267..1366

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-II human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)
VECTOR: pBAC3.5

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is AL353584; the clone sequenced to
the right is AC020674. Actual start of this clone is at base
position 1 of RP11-462H12; actual end is at base position 213793 of
RP11-462H12.

Location/Qualifiers
1..213793
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="RP11-462H12"
/clone_lib="RPCI-II"
137..322
/rpt_family="MIR"
1021..1146
/rpt_family="L2"
1267..1366
```

```

repeat_region /rpt_family="L1"
1924. .1989 /rpt_family="MIR"
repeat_region 2041. .2447 /rpt_family="ERV1"
repeat_region 2452. .2582 /rpt_family="MIR"
repeat_region 2963. .3155 /rpt_family="MIR"
repeat_region 3177. .3384 /rpt_family="L2"
repeat_region 11504. .11655 /rpt_family="MER1_type"
repeat_region 11687. .11860 /rpt_family="MIR"
repeat_region 12024. .12135 /rpt_family="MER1_type"
repeat_region 13015. .13373 /rpt_family="MALR"
repeat_region 13700. .13874 /rpt_family="MIR"
repeat_region 14678. .14876 /rpt_family="L2"
repeat_region 14892. .14966 /rpt_family="L2"
repeat_region 14971. .15126 /rpt_family="L2"
repeat_region 15743. .15852 /rpt_family="L2"
repeat_region 15927. .16094 /rpt_family="MIR"
repeat_region 16855. .16992 /rpt_family="MIR"
repeat_region 17313. .17360 /rpt_family="MER1_type"
repeat_region 17361. .17645 /rpt_family="Alu"
repeat_region 17646. .17817 /rpt_family="MER1_type"
repeat_region 18605. .18766 /rpt_family="MIR"
repeat_region 18873. .19099 /rpt_family="L2"
repeat_region 19403. .19539 /rpt_family="L2"
repeat_region 19737. .19955 /rpt_family="MIR"
repeat_region 20012. .20240 /rpt_family="MIR"
repeat_region 21006. .21161 /rpt_family="L1"
repeat_region 21473. .21845 /rpt_family="L1"
repeat_region 22083. .22284 /rpt_family="L2"
repeat_region 22473. .22630 /rpt_family="MIR"
repeat_region 23148. .23224 /rpt_family="MIR"
repeat_region 23254. .23323 /rpt_family="Mariner"
repeat_region 23345. .23414 /rpt_family="L2"
repeat_region 23417. .23568 /rpt_family="L1"
repeat_region 23924. .24290 /rpt_family="L1"
repeat_region 24291. .24880 /rpt_family="L1"
repeat_region 24970. .25140 /rpt_family="L2"
repeat_region 26009. .26156 /rpt_family="MIR"

```

```

repeat_region 26186. .26302 /rpt_family="MIR"
repeat_region 28786. .29188 /rpt_family="ERV1"
repeat_region 30166. .30520 /rpt_family="MALR"
repeat_region 30542. .31033 /rpt_family="MALR"
repeat_region 31034. .31182 /rpt_family="ERV1"
repeat_region 31183. .31487 /rpt_family="Alu"
misc_feature 31417. .31472 /note="similar to EST N71723 (NID:g1228435) yw52b09.r1"
repeat_region 31488. .31739 /rpt_family="ERV1"
repeat_region 31740. .32301 /rpt_family="ERV1"
repeat_region 32369. .32419 /rpt_family="MIR"
repeat_region 32896. .33146 /rpt_family="ERV1"
repeat_region 33437. .33489 /rpt_family="MIR"
repeat_region 33763. .34073 /rpt_family="L2"
repeat_region 35227. .35332 /rpt_family="MIR"

Alignment Scores:
Pred. No.: 9.94e+03 Length: 213793
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservatives: 1
Beat Local Similarity: 88.9% Mismatches: 0
Query Match: 93.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-7 (1-9) x AC019212 (1-213793)
QY 1 SerLeuGlnThrSerTyrValPheLeu 9
|||||
Db 82229 TCTCTGCACACTAGCTATCTATTCTG 82255

RESULT 49
AC104017
LOCUS AC104017.1 GI:17223305
DEFINITION Homo sapiens chromosome 18 clone RP11-350I3 map 18, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC104017
VERSION AC104017.1 GI:17223305
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 64140)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,B.
TITLE Homo sapiens chromosome 18, clone RP11-350I3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 64140)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,B., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,

```

Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

COMMENT

Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Suit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L21976

Center clone name: 350_I_3

* NOTE: This record contains 84 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 656: contig of 656 bp in length
 657 756: gap of 100 bp
 757 1433: contig of 677 bp in length
 1434 1533: gap of 100 bp
 1534 2210: contig of 677 bp in length
 2211 2310: gap of 100 bp
 2311 2987: contig of 677 bp in length
 2988 3087: gap of 100 bp
 3088 3741: contig of 654 bp in length
 3742 3841: gap of 100 bp
 3842 4495: contig of 654 bp in length
 4496 4595: gap of 100 bp
 4596 5252: contig of 657 bp in length
 5253 5352: gap of 100 bp
 5353 6020: contig of 668 bp in length
 6021 6120: gap of 100 bp
 6121 6854: contig of 734 bp in length
 6855 6954: gap of 100 bp
 6955 7603: contig of 649 bp in length
 7604 7703: gap of 100 bp
 7704 8384: contig of 681 bp in length
 8385 8484: gap of 100 bp
 8485 9127: contig of 643 bp in length
 9128 9227: gap of 100 bp
 9228 9908: contig of 681 bp in length
 9909 10008: gap of 100 bp
 10009 10677: contig of 669 bp in length
 10678 10777: gap of 100 bp
 10778 11450: contig of 673 bp in length
 11451 11550: gap of 100 bp
 11551 12185: contig of 635 bp in length
 12186 12885: gap of 100 bp
 12886 13016: contig of 731 bp in length
 13017 13116: gap of 100 bp
 13117 13768: contig of 652 bp in length
 13769 13868: gap of 100 bp
 13869 14532: contig of 664 bp in length

14533 14632: gap of 100 bp
 14633 15287: contig of 655 bp in length
 15288 15387: gap of 100 bp
 15388 16086: contig of 699 bp in length
 15389 16186: gap of 100 bp
 16187 16862: contig of 676 bp in length
 16863 16962: gap of 100 bp
 16964 17706: contig of 644 bp in length
 17707 17706: gap of 100 bp
 17708 18381: contig of 675 bp in length
 18382 18481: gap of 100 bp
 18483 19156: contig of 675 bp in length
 19157 19256: gap of 100 bp
 19257 19927: contig of 671 bp in length
 19928 20027: gap of 100 bp
 20028 20587: contig of 660 bp in length
 20588 20787: gap of 100 bp
 20788 21436: contig of 649 bp in length
 21437 21536: gap of 100 bp
 21537 22188: contig of 652 bp in length
 22189 22288: gap of 100 bp
 22289 22966: contig of 678 bp in length
 22967 23066: gap of 100 bp
 23068 23714: contig of 648 bp in length
 23715 23814: gap of 100 bp
 23815 24466: contig of 652 bp in length
 24467 24566: gap of 100 bp
 24567 25237: contig of 671 bp in length
 25238 25337: gap of 100 bp
 25339 25995: contig of 658 bp in length
 25996 26095: gap of 100 bp
 26096 26765: contig of 670 bp in length
 26766 26865: gap of 100 bp
 26866 27516: contig of 651 bp in length
 27517 27616: gap of 100 bp
 27617 28263: contig of 647 bp in length
 28264 28363: gap of 100 bp
 28364 29016: contig of 653 bp in length
 29017 29116: gap of 100 bp
 29117 29767: contig of 651 bp in length
 29768 29867: gap of 100 bp
 29869 30525: contig of 658 bp in length
 30526 30625: gap of 100 bp
 30626 31293: contig of 668 bp in length
 31294 31393: gap of 100 bp
 31394 32085: contig of 692 bp in length
 32086 32185: gap of 100 bp
 32186 32863: contig of 678 bp in length
 32864 32963: gap of 100 bp
 32965 33621: contig of 658 bp in length
 33622 33721: gap of 100 bp
 33722 34394: contig of 673 bp in length
 34395 34494: gap of 100 bp
 34495 35166: contig of 672 bp in length
 35167 35266: gap of 100 bp
 35267 35937: contig of 671 bp in length
 35938 36037: gap of 100 bp
 36038 36723: contig of 686 bp in length
 36724 36823: gap of 100 bp
 36824 37477: contig of 654 bp in length
 37478 37577: gap of 100 bp
 37578 38233: contig of 656 bp in length
 38234 38333: gap of 100 bp
 38334 39001: contig of 668 bp in length
 39002 39101: gap of 100 bp
 39102 39777: contig of 676 bp in length
 39778 39877: gap of 100 bp
 39879 40541: contig of 664 bp in length
 40542 40641: gap of 100 bp
 40642 41305: contig of 664 bp in length
 41306 41405: gap of 100 bp
 41406 42059: contig of 654 bp in length
 42060 42159: gap of 100 bp

* 42160 42842: contig of 683 bp in length
* 42843 42942: gap of 100 bp
* 42943 43616: contig of 674 bp in length
* 43617 43716: gap of 100 bp
* 43717 44346: contig of 630 bp in length
* 44347 44446: gap of 100 bp
* 44447 45105: contig of 659 bp in length
* 45106 45205: gap of 100 bp
* 45206 45874: contig of 669 bp in length
* 45875 45974: gap of 100 bp
* 45975 46633: contig of 659 bp in length
* 46634 46733: gap of 100 bp
* 46734 47464: contig of 731 bp in length
* 47465 47564: gap of 100 bp
* 47565 48244: contig of 680 bp in length
* 48245 48344: gap of 100 bp
* 48345 49014: contig of 670 bp in length
* 49015 49114: gap of 100 bp
* 49115 49797: contig of 683 bp in length
* 49798 49897: gap of 100 bp
* 49898 50574: contig of 677 bp in length
* 50575 50674: gap of 100 bp
* 50675 51306: contig of 632 bp in length
* 51307 51406: gap of 100 bp
* 51407 52050: contig of 644 bp in length
* 52051 52150: gap of 100 bp

Alignment Scores:

Pred. No.: 5,78e+03 Length: 64140
Score: 39.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 90.7% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-7 (1-9) x AC104017 (1-64140)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 8150 TCACACAAACAAATATGTATTCCTT 8176

RESULT 50

AP008212_089/c

WPCOMMENT

Sequence split into 308 fragments LOCUS AP008212 Accession AP008212

Fragment Name	Begin	End
AP008212_000	1	110000
AP008212_001	100001	210000
AP008212_002	200001	310000
AP008212_003	300001	410000
AP008212_004	400001	510000
AP008212_005	500001	610000
AP008212_006	600001	710000
AP008212_007	700001	810000
AP008212_008	800001	910000
AP008212_009	900001	1010000
AP008212_010	1000001	1110000
AP008212_011	1100001	1210000
AP008212_012	1200001	1310000
AP008212_013	1300001	1410000
AP008212_014	1400001	1510000
AP008212_015	1500001	1610000
AP008212_016	1600001	1710000
AP008212_017	1700001	1810000
AP008212_018	1800001	1910000
AP008212_019	1900001	2010000
AP008212_020	2000001	2110000
AP008212_021	2100001	2210000
AP008212_022	2200001	2310000
AP008212_023	2300001	2410000
AP008212_024	2400001	2510000
AP008212_025	2500001	2610000
AP008212_026	2600001	2710000

AP008212_027	2700001	2810000
AP008212_028	2800001	2910000
AP008212_029	2900001	3010000
AP008212_030	3000001	3110000
AP008212_031	3100001	3210000
AP008212_032	3200001	3310000
AP008212_033	3300001	3410000
AP008212_034	3400001	3510000
AP008212_035	3500001	3610000
AP008212_036	3600001	3710000
AP008212_037	3700001	3810000
AP008212_038	3800001	3910000
AP008212_039	3900001	4010000
AP008212_040	4000001	4110000
AP008212_041	4100001	4210000
AP008212_042	4200001	4310000
AP008212_043	4300001	4410000
AP008212_044	4400001	4510000
AP008212_045	4500001	4610000
AP008212_046	4600001	4710000
AP008212_047	4700001	4810000
AP008212_048	4800001	4910000
AP008212_049	4900001	5010000
AP008212_050	5000001	5110000
AP008212_051	5100001	5210000
AP008212_052	5200001	5310000
AP008212_053	5300001	5410000
AP008212_054	5400001	5510000
AP008212_055	5500001	5610000
AP008212_056	5600001	5710000
AP008212_057	5700001	5810000
AP008212_058	5800001	5910000
AP008212_059	5900001	6010000
AP008212_060	6000001	6110000
AP008212_061	6100001	6210000
AP008212_062	6200001	6310000
AP008212_063	6300001	6410000
AP008212_064	6400001	6510000
AP008212_065	6500001	6610000
AP008212_066	6600001	6710000
AP008212_067	6700001	6810000
AP008212_068	6800001	6910000
AP008212_069	6900001	7010000
AP008212_070	7000001	7110000
AP008212_071	7100001	7210000
AP008212_072	7200001	7310000
AP008212_073	7300001	7410000
AP008212_074	7400001	7510000
AP008212_075	7500001	7610000
AP008212_076	7600001	7710000
AP008212_077	7700001	7810000
AP008212_078	7800001	7910000
AP008212_079	7900001	8010000
AP008212_080	8000001	8110000
AP008212_081	8100001	8210000
AP008212_082	8200001	8310000
AP008212_083	8300001	8410000
AP008212_084	8400001	8510000
AP008212_085	8500001	8610000
AP008212_086	8600001	8710000
AP008212_087	8700001	8810000
AP008212_088	8800001	8910000
AP008212_089	8900001	9010000
AP008212_090	9000001	9110000
AP008212_091	9100001	9210000
AP008212_092	9200001	9310000
AP008212_093	9300001	9410000
AP008212_094	9400001	9510000
AP008212_095	9500001	9610000
AP008212_096	9600001	9710000
AP008212_097	9700001	9810000
AP008212_098	9800001	9910000
AP008212_099	9900001	10010000

AP008212_100 10000001 10110000
AP008212_101 10100001 10210000
AP008212_102 10200001 10310000
AP008212_103 10300001 10410000
AP008212_104 10400001 10510000
AP008212_105 10500001 10610000
AP008212_106 10600001 10710000
AP008212_107 10700001 10810000
AP008212_108 10800001 10910000
AP008212_109 10900001 11010000
AP008212_110 11000001 11110000
AP008212_111 11100001 11210000
AP008212_112 11200001 11310000
AP008212_113 11300001 11410000
AP008212_114 11400001 11510000
AP008212_115 11500001 11610000
AP008212_116 11600001 11710000
AP008212_117 11700001 11810000
AP008212_118 11800001 11910000
AP008212_119 11900001 12010000
AP008212_120 12000001 12110000
AP008212_121 12100001 12210000
AP008212_122 12200001 12310000
AP008212_123 12300001 12410000
AP008212_124 12400001 12510000
AP008212_125 12500001 12610000
AP008212_126 12600001 12710000
AP008212_127 12700001 12810000
AP008212_128 12800001 12910000
AP008212_129 12900001 13010000
AP008212_130 13000001 13110000
AP008212_131 13100001 13210000
AP008212_132 13200001 13310000
AP008212_133 13300001 13410000
AP008212_134 13400001 13510000
AP008212_135 13500001 13610000
AP008212_136 13600001 13710000
AP008212_137 13700001 13810000
AP008212_138 13800001 13910000
AP008212_139 13900001 14010000
AP008212_140 14000001 14110000
AP008212_141 14100001 14210000
AP008212_142 14200001 14310000
AP008212_143 14300001 14410000
AP008212_144 14400001 14510000
AP008212_145 14500001 14610000
AP008212_146 14600001 14710000
AP008212_147 14700001 14810000
AP008212_148 14800001 14910000
AP008212_149 14900001 15010000
AP008212_150 15000001 15110000
AP008212_151 15100001 15210000
AP008212_152 15200001 15310000
AP008212_153 15300001 15410000
AP008212_154 15400001 15510000
AP008212_155 15500001 15610000
AP008212_156 15600001 15710000
AP008212_157 15700001 15810000
AP008212_158 15800001 15910000
AP008212_159 15900001 16010000
AP008212_160 16000001 16110000
AP008212_161 16100001 16210000
AP008212_162 16200001 16310000
AP008212_163 16300001 16410000
AP008212_164 16400001 16510000
AP008212_165 16500001 16610000
AP008212_166 16600001 16710000
AP008212_167 16700001 16810000
AP008212_168 16800001 16910000
AP008212_169 16900001 17010000
AP008212_170 17000001 17110000
AP008212_171 17100001 17210000
AP008212_172 17200001 17310000

AP008212_173 17300001 17410000
AP008212_174 17400001 17510000
AP008212_175 17500001 17610000
AP008212_176 17600001 17710000
AP008212_177 17700001 17810000
AP008212_178 17800001 17910000
AP008212_179 17900001 18010000
AP008212_180 18000001 18110000
AP008212_181 18100001 18210000
AP008212_182 18200001 18310000
AP008212_183 18300001 18410000
AP008212_184 18400001 18510000
AP008212_185 18500001 18610000
AP008212_186 18600001 18710000
AP008212_187 18700001 18810000
AP008212_188 18800001 18910000
AP008212_189 18900001 19010000
AP008212_190 19000001 19110000
AP008212_191 19100001 19210000
AP008212_192 19200001 19310000
AP008212_193 19300001 19410000
AP008212_194 19400001 19510000
AP008212_195 19500001 19610000
AP008212_196 19600001 19710000

Alignment Scores:

Pred. No.: 9.12e+03 Length: 110000
Score: 39.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.7% Indels: 0
DB: 15 Gaps: 0

US-10-774-176-7 (1-9) x AP008212_089 (1-110000)

Qy 2 LeuGlnThrSerTyrValPheLeu 9

Db 15127 TTGCARACTAGTTAIGTTTCCTA 15104

Search completed: April 25, 2006, 20:34:54

Job time : 3088.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:56:14 ; Search time 295.3 Seconds

(without alignments)
203.123 Million cell updates/sec

Title: US-10-774-176-6

Perfect score: 40

Sequence: 1 ALIGAIFLL 9

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abse/ABSSWB spool/US10774176/runat 24042006 165112 19185/app query.fasta 1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNIT8-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=1000
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abse805p
-USER=US10774176 @CGN 1 1 3463 @runat 24042006 165112 19185 -NCPU=6 -ICPU=3
-NO MMAP -NRG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 21.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	246	10	ADK11641 Breast ca
2	40	100.0	475	13	ADU11677 Solid tum
3	40	100.0	901	3	AAA27060 Canine 5T
4	40	100.0	927	6	ABT07721 Breast ca

5	40	100.0	927	8	ABX76333
6	40	100.0	927	10	ADB80503
7	40	100.0	927	11	ADN38723
8	40	100.0	973	8	AAU56198
9	40	100.0	1156	6	ABV99349
10	40	100.0	1260	6	ABK87175
11	40	100.0	1260	10	ADB97513
12	40	100.0	1260	10	ADB97452
13	40	100.0	1263	3	AAA27058
14	40	100.0	1263	4	AAF89736
15	40	100.0	1263	6	ABK87174
16	40	100.0	1281	3	AAA27059
17	40	100.0	1331	8	AAU56199
18	40	100.0	2020	10	ADJ56299
19	40	100.0	2053	8	ACC51052
20	40	100.0	2053	8	ABX76332
21	40	100.0	2053	8	AAU56197
22	40	100.0	2053	8	AAU56200
23	40	100.0	2053	11	ADN38721
24	40	100.0	2053	12	ADL06473
25	40	100.0	2053	12	ADN03961
26	40	100.0	2053	13	ADR25444
27	40	100.0	2053	13	ACN38510
28	40	100.0	2053	13	ADV35098
29	40	100.0	2338	5	AA887175
30	40	100.0	2359	4	AAK94253
31	40	100.0	2359	12	ADL30831
32	40	100.0	2361	4	AAK94254
33	40	100.0	2361	12	ADL26162
34	40	100.0	2361	12	ADL30833
35	40	100.0	2361	12	ADL26160
36	40	100.0	2557	12	ADL26158
37	39	97.5	1494	10	ADF02237
38	38	95.0	433	6	ABN60581
39	38	95.0	567	6	ABN60925
40	38	95.0	1101	12	ADQ35702
41	37	92.5	951	8	ACA28481
42	36	90.0	225	6	ABN70062
43	36	90.0	225	6	ABN69887
44	36	90.0	453	5	AA887174
45	36	90.0	681	10	ACF71597
46	36	90.0	687	10	ADP03189
47	36	90.0	922	9	AAU57769
48	36	90.0	922	9	AAU57769
49	36	90.0	3222	4	ABL05253
50	36	90.0	8737	4	ABL05252
51	36	90.0	110000	10	ACF67367_51
52	36	90.0	110000	10	ACF67367_52
53	36	90.0	110000	10	ACF65387_3
54	36	90.0	349980	5	AAF86431
55	35	87.5	231	13	ADU02171
56	35	87.5	455	3	AA243023
57	35	87.5	455	13	ADU72185
58	35	87.5	455	14	ADZ73176
59	35	87.5	490	4	AAF67953
60	35	87.5	494	2	AAU19491
61	35	87.5	603	5	AAH66087
62	35	87.5	624	8	AB254726
63	35	87.5	644	13	ADR25740
64	35	87.5	654	2	AAZ30614
65	35	87.5	654	3	AAA72905
66	35	87.5	672	5	AAH66086
67	35	87.5	684	5	AA831338
68	35	87.5	750	5	ADM19652
69	35	87.5	752	5	ADM19412
70	35	87.5	836	12	ADL12884
71	35	87.5	1152	12	ADO20144
72	35	87.5	1152	13	ADP25146
73	35	87.5	1152	14	ADY17683
74	35	87.5	1288	4	AA67951
75	35	87.5	1314	8	ACA37128
76	35	87.5	1341	8	ACA23707
77	35	87.5	1464	10	ACF70187

Abx76333	Lung canc
Adb80503	Ovarian c
Adn38723	Cancer/an
Adv99349	Human LRR
Abk87175	cdNA enco
Adb97513	Peline 5T
Adb97452	DNA enco
Aaa27058	Human 5T4
Aaf89736	Nucleotid
Abk87174	cdNA enco
Aaa27059	Mouse 5T4
Ad56199	Human LRR
Adj56299	Human CDN
Acc51052	Human bla
Abx76332	Lung canc
Ad56197	Human LRR
Ad56200	Human LRR
Adn38721	Cancer/an
Adl06473	Human tum
Adn03961	Antipsori
Adr25444	Breast ca
Acn38510	Tumour-as
Adv35098	Human cdn
Aa887175	DNA enco
Aak94253	Human ful
Adl30831	Full leng
Aak94254	Human ful
Adl30833	Full leng
Adl26160	Human cdn
Adl26158	Human cdn
Adf02237	Bacterial
Abn60581	Human can
Abn60925	Human can
Ad35702	Novel mou
ACA28481	Prokaryot
ABN70062	Streptoco
ABN69887	Streptoco
Aa887174	DNA enco
ACF71597	Photorhab
ADP03189	Bacterial
AAU57769	Green flu
AAU57769	Aequorea
ABL05253	Drosophil
ABL05252	Drosophil
Continuation	(52 o
Continuation	(4 of
Aaf86431	Pyrococcu
ADU02171	Novel hum
Aaz43023	Human 5'
ADU72185	Signal pe
ADZ73176	Human 5'
Aaf67953	Corynebac
Aat19491	Human gen
Aah66087	C glutami
Abz54726	Aspergill
ADR25740	Breast ca
Aaz30614	Human glu
Aaa72905	Human glu
Aah66086	C glutami
AA831338	DNA enco
ADM19652	Novel hum
ADM19412	Novel hum
ADL12884	Human sce
ADO20144	Human PRO
ADP25146	PRO polyP
ADY17683	DNA enco
Aaf67951	Corynebac
ACA37128	Prokaryot
ACA23707	Prokaryot
ACF70187	Photorhab

78	35	87.5	1558	3	AAF12731	151	33	82.5	1152	11	ABD03187
79	35	87.5	1558	13	ADU56772	152	33	82.5	1152	11	AAF67985
80	35	87.5	1558	14	ADU294775	153	33	82.5	1152	9	ADA32154
81	35	87.5	2009	10	ADA44864	154	33	82.5	1902	2	AAV52612
82	35	87.5	2009	10	ADA44860	155	33	82.5	1902	2	AAV52611
83	35	87.5	2148	2	AAV52612	156	33	82.5	2000	10	ACC60867
84	35	87.5	2435	14	ABE68330	157	33	82.5	2000	10	ADK62247
85	35	87.5	2809	14	ABE68330	158	33	82.5	2181	11	ACL28439
86	35	87.5	3130	14	ABE68328	159	33	82.5	2578	12	ADH51271
87	35	87.5	3428	10	ACC72803	160	33	82.5	2819	4	ABL23110
88	35	87.5	3428	10	ACC72805	161	33	82.5	3009	4	ABL26771
89	35	87.5	4069	12	ADQ24787	162	33	82.5	3993	13	ADU01816
90	35	87.5	6422	4	AAV52612	163	33	82.5	5293	4	ABL26770
91	35	87.5	6422	8	ACF64584	164	33	82.5	5998	2	AAV52612
92	35	87.5	4085	10	ADU74277	165	33	82.5	5998	6	ABS98851
93	35	87.5	110000	6	ABX09336	166	33	82.5	7480	14	ADY72849
94	35	87.5	110000	10	ACF65385	167	33	82.5	7574	2	AAV52612
95	35	87.5	110000	10	ACF67367	168	33	82.5	7574	2	ABS98884
96	35	87.5	110000	12	ADJ25985	169	33	82.5	10509	13	ADT05511
97	35	87.5	110000	12	ADN97989	170	33	82.5	12660	6	ADH48717
98	35	87.5	110000	12	ADN97989	171	33	82.5	13686	5	AAV52612
99	35	87.5	110000	12	ADN97989	172	33	82.5	30274	5	AAV52612
100	35	87.5	110000	14	ABE39174	173	33	82.5	53809	10	ADC86800
101	35	87.5	110000	14	ABE39174	174	33	82.5	72352	12	ADQ97067
102	35	87.5	110000	14	ABE39174	175	33	82.5	72352	12	ADQ97067
103	35	87.5	110000	14	ABE39174	176	33	82.5	72352	12	ADQ97067
104	35	87.5	110000	14	ABE39174	177	33	82.5	81684	11	ACN43944
105	35	87.5	110000	14	ABE39174	178	33	82.5	110000	2	AAV52612
106	35	87.5	110000	14	ABE39174	179	33	82.5	110000	2	AAV52612
107	35	87.5	34980	5	AAH68527	180	33	82.5	110000	6	ABQ69245
108	34	85.0	361	12	ADP94590	181	33	82.5	110000	6	ABQ69245
109	34	85.0	645	4	AAK82352	182	33	82.5	110000	6	ABQ69245
110	34	85.0	849	8	ACA54049	183	33	82.5	110000	13	ABD32806
111	34	85.0	904	2	AAV06139	184	33	82.5	110000	13	ABD32806
112	34	85.0	904	12	ADU74145	185	33	82.5	110000	13	ABD32806
113	34	85.0	1048	8	AAV52612	186	33	82.5	128361	14	ADX07563
114	34	85.0	1089	5	AAV52612	187	33	82.5	128361	14	ADX07563
115	34	85.0	1258	8	ABX05055	188	33	82.5	130263	6	ABK83573
116	34	85.0	1428	8	ACA47696	189	33	82.5	200418	11	ACN44226
117	34	85.0	2247	6	ABZ13286	190	33	82.5	349980	5	AAH68529
118	34	85.0	2370	13	ADT1641	191	33	82.5	349980	5	AAH68529
119	34	85.0	4467	2	AAV52612	192	33	82.5	349980	5	AAH68529
120	34	85.0	30753	14	ABE61114	193	33	82.5	349980	5	AAH68529
121	34	85.0	49999	2	AAV52612	194	33	82.5	349980	5	AAH68529
122	34	85.0	49999	2	AAV52612	195	33	82.5	349980	5	AAH68529
123	34	85.0	241748	14	ADZ13116	196	33	82.5	349980	5	AAH68529
124	33	82.5	94	12	ADG99473	197	33	82.5	349980	5	AAH68529
125	33	82.5	213	10	ADH84227	198	33	82.5	349980	5	AAH68529
126	33	82.5	243	8	ACA45542	199	33	82.5	349980	5	AAH68529
127	33	82.5	273	6	ABK44688	200	33	82.5	349980	5	AAH68529
128	33	82.5	294	10	ADC93622	201	33	82.5	349980	5	AAH68529
129	33	82.5	407	4	ABA08942	202	33	82.5	349980	5	AAH68529
130	33	82.5	407	4	AAK53329	203	33	82.5	349980	5	AAH68529
131	33	82.5	531	10	ADH84226	204	33	82.5	349980	5	AAH68529
132	33	82.5	539	13	ACN46694	205	33	82.5	349980	5	AAH68529
133	33	82.5	587	10	ADG57091	206	33	82.5	349980	5	AAH68529
134	33	82.5	612	10	ADG37557	207	33	82.5	349980	5	AAH68529
135	33	82.5	612	10	ADG37743	208	33	82.5	349980	5	AAH68529
136	33	82.5	693	11	ABD03572	209	33	82.5	349980	5	AAH68529
137	33	82.5	744	4	AAV67987	210	33	82.5	349980	5	AAH68529
138	33	82.5	762	11	ABD14320	211	33	82.5	349980	5	AAH68529
139	33	82.5	819	4	ABL23111	212	33	82.5	349980	5	AAH68529
140	33	82.5	880	3	AAK30856	213	33	82.5	349980	5	AAH68529
141	33	82.5	882	6	ABQ92505	214	33	82.5	349980	5	AAH68529
142	33	82.5	969	8	ACA30522	215	33	82.5	349980	5	AAH68529
143	33	82.5	993	8	ACA26128	216	33	82.5	349980	5	AAH68529
144	33	82.5	1011	11	ABD13956	217	33	82.5	349980	5	AAH68529
145	33	82.5	1032	4	AAV54152	218	33	82.5	349980	5	AAH68529
146	33	82.5	1074	5	AAH66760	219	33	82.5	349980	5	AAH68529
147	33	82.5	1100	3	AAK30855	220	33	82.5	349980	5	AAH68529
148	33	82.5	1110	11	ABD03254	221	33	82.5	349980	5	AAH68529
149	33	82.5	1116	11	ABD03500	222	33	82.5	349980	5	AAH68529
150	33	82.5	1149	6	ABR90338	223	33	82.5	349980	5	AAH68529

224	32	80.0	663	10	ADH82701	Adh82701 Enterococ	297	32	80.0	1890	8	ACF74345	Acf74345 Staphyloc
225	32	80.0	705	6	ABQ82699	Abq82699 Mouse imm	c 298	32	80.0	2045	12	ADJ62780	Adj62780 Human cDN
226	32	80.0	711	13	ADX14872	Adx14872 Plant ful	c 299	32	80.0	2045	14	ADX07279	Adx07279 Cyclin-de
227	32	80.0	800	13	ADR63309	Adr63309 Cotton cd	c 300	32	80.0	2098	13	ADX14102	Adx14102 Plant ful
228	32	80.0	803	4	AAF22579	Aaf22579 Human bre	301	32	80.0	2173	10	ADB62577	Adb62577 Human cDN
229	32	80.0	816	14	ADY77510	Ady77510 A. thalia	302	32	80.0	2315	10	ADB63664	Adb63664 Human cDN
230	32	80.0	833	4	AAF22581	Aaf22581 Human bre	c 303	32	80.0	2483	10	ADB53171	Adb53171 Primary r
231	32	80.0	939	6	ABK75647	Abk75647 Bacillus	c 304	32	80.0	2483	13	ADV41350	Adv41350 Rat card
232	32	80.0	948	4	AAH32300	Aah32300 Human oif	c 305	32	80.0	2537	4	AAH14602	Aah14602 Human cDN
233	32	80.0	948	4	AAH31972	Aah31972 Human oif	c 306	32	80.0	2554	6	ABA92314	Aba92314 Human cyl
234	32	80.0	951	5	AAH42234	Aah42234 Human cDN	307	32	80.0	2598	4	AAH13924	Aah13924 Human cDN
235	32	80.0	951	6	ABK16630	Abk16630 Human G-c	308	32	80.0	2601	2	AAV01537	Aav01537 Acylcoenz
236	32	80.0	951	6	ABK68540	Abk68540 Human DNA	309	32	80.0	2601	2	AAV96369	Aav96369 Yeast acy
237	32	80.0	951	6	ABK37522	Abk37522 DNA encod	c 310	32	80.0	2637	6	ABO76648	Abg76648 C. albica
238	32	80.0	961	10	ADD12728	Add12728 cDNA enco	c 311	32	80.0	2637	6	ABZ32228	Abz32228 Candida a
239	32	80.0	962	6	AAF88403	Aaf88403 Human GPC	c 312	32	80.0	2826	2	AAV22225	Aav22225 SIRE-1 ge
240	32	80.0	962	12	ACH87662	Ach87662 Human gen	c 313	32	80.0	2826	10	ADG28403	Adg28403 Soybean S
241	32	80.0	1003	3	AAF22509	Aaf22509 Human bre	314	32	80.0	2827	12	ADK70373	Adk70373 Respirato
242	32	80.0	1190	3	AAC51552	Aac51552 Arabidops	315	32	80.0	2880	4	AAF77095	Aaf77095 Arabidops
243	32	80.0	1192	3	AAC39286	Aac39286 Arabidops	316	32	80.0	2880	10	ACF36550	Acf36550 Arabidops
244	32	80.0	1199	3	AAC49104	Aac49104 Arabidops	c 317	32	80.0	3055	10	ACA56858	Aca56858 Human sig
245	32	80.0	1203	3	AAC40238	Aac40238 Arabidops	c 318	32	80.0	3055	12	ADI56654	Adi56654 Human pol
246	32	80.0	1242	13	ADU25709	Adu25709 cDNA enco	c 319	32	80.0	3122	12	ADP21378	Adp21378 Gene P2RY
247	32	80.0	1251	6	AAH96327	Aah96327 Arabidops	c 320	32	80.0	3122	14	ABA50186	Aba50186 P2RY1 cod
248	32	80.0	1254	6	ABZ17562	Abz17562 Arabidops	c 321	32	80.0	3338	4	ABL06617	Abi06617 Drosophil
249	32	80.0	1279	14	ACL67915	ACL67915 M. xanthu	c 322	32	80.0	3349	4	ABL19617	Abi19617 Drosophil
250	32	80.0	1280	10	ADB31439	Adb31439 Bicalutam	c 323	32	80.0	3564	4	ABL19615	Abi19615 Drosophil
251	32	80.0	1280	14	ADZ75514	Adz75514 Human EST	324	32	80.0	3603	14	AEA19628	Aea19628 Novel hum
252	32	80.0	1301	7	ADG31018	Adg31018 Human gen	325	32	80.0	3617	4	AAH24799	Aah24799 Nucleotid
253	32	80.0	1301	7	ADY36406	Ady36406 HTRA geno	c 326	32	80.0	3639	4	ABA88913	Abg88913 Escherich
254	32	80.0	1306	13	ADT44719	Adt44719 Bacterial	c 327	32	80.0	3754	4	ABL19619	Abi19619 Drosophil
255	32	80.0	1314	4	AAH33296	Aah33296 DNA encod	328	32	80.0	3778	4	ABL14988	Abi14988 Drosophil
256	32	80.0	1314	4	AAH34943	Aah34943 cDNA enco	c 329	32	80.0	3806	13	ADR08439	Adr08439 Full leng
257	32	80.0	1314	10	ADC46101	Adc46101 Human neo	c 330	32	80.0	4291	14	ABZ27085	Abz27085 Pinus rad
258	32	80.0	1332	13	ADT19631	Adt19631 Plant cDN	c 331	32	80.0	4336	5	AAO05447	Aao05447 Human sec
259	32	80.0	1351	10	ADC85952	Adc85952 Human GPC	c 332	32	80.0	4764	9	ACD27526	Acd27526 Human cDN
260	32	80.0	1368	10	ADI21915	Adi21915 Novel hum	333	32	80.0	4764	10	ADG32900	Adg32900 Human DNA
261	32	80.0	1368	11	ACH99857	Ach99857 Klebsiell	334	32	80.0	4764	12	ADQ17463	Adq17463 Human sof
262	32	80.0	1392	8	ACA54092	Aca54092 Prokaryot	335	32	80.0	4764	12	ADQ96674	Adq96674 Human rib
263	32	80.0	1415	3	AAC53142	Aac53142 Arabidops	336	32	80.0	4764	13	ADR68726	Adr68726 Human cDN
264	32	80.0	1415	3	AAC54854	Aac54854 Arabidops	337	32	80.0	4764	14	ADX03768	Adx03768 Human cDN
265	32	80.0	1442	2	AAH13632	Aah13632 Enterococ	338	32	80.0	4764	14	ADY17385	Ady17385 DNA encod
266	32	80.0	1442	6	ABS99427	Abs99427 Enterococ	339	32	80.0	4764	14	ABE47410	Abe47410 Human rib
267	32	80.0	1455	4	AAH33094	Aah33094 DNA encod	340	32	80.0	4971	4	ABL29542	Abi29542 Drosophil
268	32	80.0	1458	4	AAH34779	Aah34779 cDNA enco	341	32	80.0	5170	8	ABZ58595	Abz58595 Green flu
269	32	80.0	1458	10	ADC45937	Adc45937 Human neo	342	32	80.0	5170	12	ADI10301	Adi10301 Asquorea
270	32	80.0	1467	4	AAH53033	Aah53033 Enterococ	343	32	80.0	5170	12	ADO24330	Ado24330 Asquorea
271	32	80.0	1467	4	ABA82964	Aba82964 Enterococ	c 344	32	80.0	5400	10	ADH83888	Adh83888 Enterococ
272	32	80.0	1496	10	ADD30415	Add30415 Plant yie	345	32	80.0	5425	12	ADQ22183	Adq22183 Human sof
273	32	80.0	1496	14	ADZ00616	Adz00616 G2157 pol	c 346	32	80.0	5586	4	AAK69027	Aak69027 Human imm
274	32	80.0	1542	6	ABK35258	Abk35258 Human cDN	c 347	32	80.0	6035	2	AAV74583	Aav74583 Staphyloc
275	32	80.0	1549	10	ADI15977	Adi15977 Human PP	348	32	80.0	6050	10	ADE84025	Ade84025 5' regula
276	32	80.0	1549	12	ADJ81703	Adj81703 Tumour an	349	32	80.0	6121	6	ABL33974	Abi33974 Human imm
277	32	80.0	1609	8	ACA32118	Aca32118 Prokaryot	350	32	80.0	6122	13	ABD32871	Abd32871 Mouse can
278	32	80.0	1629	9	ACF25385	Acf25385 Human put	c 351	32	80.0	6234	4	AAI98984	Aai98984 Human exc
279	32	80.0	1629	4	AAH51980	Aah51980 Mycobacte	c 352	32	80.0	6234	5	AAI63334	Aai63334 Human kid
280	32	80.0	1629	8	ACA35704	Aca35704 Prokaryot	353	32	80.0	6258	13	ABD32874	Abd32874 Human can
281	32	80.0	1647	11	ACH95536	Ach95536 Klebsiell	c 354	32	80.0	6721	6	AAH18600	Aah18600 Purinergi
282	32	80.0	1656	11	ACH96117	Ach96117 Klebsiell	c 355	32	80.0	7037	11	AAH18599	Aah18599 Purinergi
283	32	80.0	1679	13	ADX48398	Adx48398 Plant ful	356	32	80.0	7037	13	ACN44891	Acn44891 Human mRN
284	32	80.0	1683	4	AAH87124	Aah87124 NOV13 cod	357	32	80.0	7039	13	ABD32875	Abd32875 Human can
285	32	80.0	1683	6	ABK11586	Abk11586 S. agalac	358	32	80.0	7053	2	AAH13015	Aah13015 Enterococ
286	32	80.0	1692	5	AAO02391	Aao02391 Virulent	359	32	80.0	7053	6	ABH98810	Abh98810 Enterococ
287	32	80.0	1725	6	ABQ70748	Abq70748 Listeria	360	32	80.0	7296	2	AAH12996	Aah12996 Enterococ
288	32	80.0	1765	10	ADJ56437	Adj56437 Human cDN	361	32	80.0	7296	6	ABS98791	Abs98791 Enterococ
289	32	80.0	1768	4	ABL14989	Abi14989 Drosophil	362	32	80.0	7313	13	ACN42227	Acn42227 Human dia
290	32	80.0	1784	14	ABA20644	Aba20644 Novel hum	363	32	80.0	7618	14	ADK08027	Adk08027 Cyclin-de
291	32	80.0	1788	4	ABA88917	Abg88917 Escherich	364	32	80.0	7618	14	ADZ49064	Adz49064 Inulin s
292	32	80.0	1794	4	ABL29543	Abi29543 Drosophil	365	32	80.0	7926	13	ACN42226	Acn42226 Human dia
293	32	80.0	1833	4	AAH27155	Aah27155 Yeast ARE	c 366	32	80.0	8160	2	AAH13096	Aah13096 Enterococ
294	32	80.0	1833	10	ADK65658	Adk65658 S cerevis	c 367	32	80.0	8160	6	ABS98891	Abg98891 Enterococ
295	32	80.0	1833	13	ADS46786	Adg46786 Bacterial	c 368	32	80.0	8382	4	ABL04086	Abi04086 Drosophil
296	32	80.0	1847	11	ACN44017	Acn44017 Mouse mRN	c 369	32	80.0	9072	10	ADG28440	Adg28440 Soybean S

C 370	32	80.0	9358	10	ADG28443	Adg28443 Soybean S	443	31	77.5	183	4	AAH46910	Aah46910 cDNA enco
C 371	32	80.0	9399	10	ADG28446	Adg28446 Soybean S	444	31	77.5	198	4	AAH48603	Aah48603 Pseudomon
C 372	32	80.0	10412	12	ADJ81641	Adj81641 Human tyr	445	31	77.5	198	4	AAH48609	Aah48609 Pseudomon
C 373	32	80.0	10663	3	AAAL2624	Aaa12624 Genomic D	446	31	77.5	198	8	ACA15654	AcA15654 Prokaryot
C 374	32	80.0	12050	12	ADN29130	Adn29130 Human kal	447	31	77.5	198	8	ACA15655	AcA15655 Prokaryot
C 375	32	80.0	12483	13	ADR72626	Adr72626 Human ren	c 448	31	77.5	202	4	AAK54633	Aak54633 Human hae
C 376	32	80.0	12483	13	ADR72878	Adr72878 Human ova	c 449	31	77.5	202	4	AAK54758	Aak54758 Human hae
C 377	32	80.0	12483	13	ADY67596	Ady67596 Human kal	450	31	77.5	205	4	AAH48558	Aah48558 Pseudomon
C 378	32	80.0	14066	6	ABR78923	AbR78923 E. coli C	451	31	77.5	205	4	AAH48597	Aah48597 Pseudomon
C 379	32	80.0	14066	10	ADH80490	Adh80490 Escherich	452	31	77.5	205	8	ACA15604	AcA15604 Prokaryot
C 380	32	80.0	15424	4	ABL29522	AbI29522 Drosophi	453	31	77.5	205	8	ACA15647	AcA15647 Prokaryot
C 381	32	80.0	20303	2	AAAT71699	Aat71699 Human deo	c 454	31	77.5	209	5	ABV59345	Abv59345 Human pro
C 382	32	80.0	21469	4	AAK89568	Aak89568 Human dig	c 455	31	77.5	211	5	ABV59361	Abv59361 Human pro
C 383	32	80.0	21475	4	AAK89569	Aak89569 Human dig	456	31	77.5	219	6	ABN75121	Abn75121 Human tra
C 384	32	80.0	22477	11	ACN44910	Acn44910 Human gen	c 457	31	77.5	228	6	ABN77757	Abn77757 Human ORF
C 385	32	80.0	25320	11	ACN44016	Acn44016 Mouse gen	458	31	77.5	243	8	ACA43697	AcA43697 Prokaryot
C 386	32	80.0	26278	14	AEb32413	Aeb32413 Human gen	c 459	31	77.5	253	4	AAK54637	Aak54637 Human hae
C 387	32	80.0	26764	2	AAAT71696	Aat71696 Human deo	c 460	31	77.5	253	4	AAK54761	Aak54761 Human hae
C 388	32	80.0	29272	14	ACL64754	Acl64754 M. xanthu	461	31	77.5	264	6	ABQ67307	Abq67307 Listeria
C 389	32	80.0	29729	2	AAI13175	Aai13175 Enterococ	462	31	77.5	264	6	ABQ69555	Abq69555 Listeria
C 390	32	80.0	29729	6	ABR98970	AbR98970 Enterococ	c 463	31	77.5	282	4	AAK69067	Aak69067 Soybean K
C 391	32	80.0	29871	6	ABR86359	AbR86359 L. lactis	c 464	31	77.5	300	2	AAZ14285	Aaz14285 Human gen
C 392	32	80.0	34604	14	AEb32362	Aeb32362 Human gen	c 465	31	77.5	313	4	AAI12133	Aai12133 Probe #20
C 393	32	80.0	43419	10	ADC86998	Adc86998 Human GPC	c 466	31	77.5	313	4	ABA53839	AbA53839 Human foe
C 394	32	80.0	48680	11	ACN45210	Acn45210 Human gen	c 467	31	77.5	313	4	AAI33477	Aai33477 Probe #21
C 395	32	80.0	50600	4	AAK85994	Aak85994 Human imm	c 468	31	77.5	313	4	ABA43392	AbA43392 Human bre
C 396	32	80.0	71850	12	ADQ97478	Adq97478 Mouse can	c 469	31	77.5	313	4	ABR23583	AbR23583 Probe #20
C 397	32	80.0	90435	12	ADQ59524	Adq59524 Human can	c 470	31	77.5	313	4	AAK27550	Aak27550 Human bon
C 398	32	80.0	90537	14	ADZ13905	Adz13905 Human can	c 471	31	77.5	313	4	AAK02098	Aak02098 Human bra
C 399	32	80.0	104451	13	ABD33121	Abd33121 Human can	c 472	31	77.5	313	4	ABS27118	AbS27118 Human liv
C 400	32	80.0	110000	2	AAZ01425_09	Continuation (10 o	c 473	31	77.5	313	5	AAI02048	Aai02048 Probe #20
C 401	32	80.0	110000	4	AAK91990_04	Continuation (5 of	c 474	31	77.5	316	8	ABS02028	AbS02028 Human gen
C 402	32	80.0	110000	4	AAK95240_00	Aak95240 Human neu	c 475	31	77.5	340	8	ABX48197	Abx48197 Bovine ES
C 403	32	80.0	110000	4	AAI199682_09	Continuation (10 o	c 476	31	77.5	340	4	AAK54629	Aak54629 Human hae
C 404	32	80.0	110000	4	AAI199683_09	Continuation (10 o	c 477	31	77.5	340	4	AAK54673	Aak54673 Human hae
C 405	32	80.0	110000	6	AAK96733_00	Aak96733 Human neu	c 478	31	77.5	340	4	AAK54669	Aak54669 Human hae
C 406	32	80.0	110000	6	ABT00010_00	Abt00010 Human neu	c 479	31	77.5	340	4	AAK54779	Aak54779 Human hae
C 407	32	80.0	110000	6	ABQ69245_22	Continuation (23 o	c 480	31	77.5	340	4	AAK54756	Aak54756 Human hae
C 408	32	80.0	110000	6	ABT01503_00	Abt01503 Human neu	c 481	31	77.5	340	4	AAK54775	Aak54775 Human hae
C 409	32	80.0	110000	6	ABAO3041_22	Continuation (23 o	c 482	31	77.5	340	4	AAK54748	Aak54748 Human hae
C 410	32	80.0	110000	9	ACH03408_1	Continuation (2 of	c 483	31	77.5	340	4	AAK54600	Aak54600 Human hae
C 411	32	80.0	110000	10	ACF67367_14	Continuation (15 o	c 484	31	77.5	340	10	ADD67587	Add67587 Human CD7
C 412	32	80.0	110000	10	ACF65387_2	Continuation (3 of	c 485	31	77.5	340	10	ADD67585	Add67585 Human CD7
C 413	32	80.0	110000	11	ADW70291_00	Adw70291 Human neu	c 486	31	77.5	374	8	ABX54868	Abx54868 Bovine BS
C 414	32	80.0	110000	12	ADH77486_00	Adh77486 Human neu	c 487	31	77.5	384	5	AAH79337	Aah79337 DNA encod
C 415	32	80.0	110000	12	ADK16049_2	Continuation (3 of	c 488	31	77.5	385	4	ABS30308	AbS30308 Human liv
C 416	32	80.0	110000	12	ADQ34435_5	Continuation (6 of	c 489	31	77.5	415	4	ABL27251	AbI27251 Drosophi
C 417	32	80.0	111331	13	ABD32870_1	Abd32870 Mouse can	c 490	31	77.5	417	4	AAI192617	Aai192617 Human pol
C 418	32	80.0	119235	12	ADP03020	Adp03020 Human hou	c 491	31	77.5	419	5	ABV16870	Abv16870 Human pro
C 419	32	80.0	119241	13	ADH88518	Adh88518 Human hou	c 492	31	77.5	452	4	AAH35457	Aah35457 Human col
C 420	32	80.0	130241	13	ADU60136	Adu60136 Housekeep	c 493	31	77.5	453	12	ADP66212	Adp66212 Human CDN
C 421	32	80.0	130244	13	ABD32872	Abd32872 Human can	c 494	31	77.5	469	5	ABV46667	Abv46667 Human pro
C 422	32	80.0	131576	11	ACN44890	Acn44890 Human gen	c 495	31	77.5	469	6	ABQ93891	Abq93891 Human epi
C 423	32	80.0	131680	10	ADF29092	Adf29092 Agrotis s	c 496	31	77.5	469	12	ADO10064	Ado10064 Novel hum
C 424	32	80.0	139308	8	ABE12769	AbE12769 Human PRK	c 497	31	77.5	471	6	ABN68414	Abn68414 Streptoco
C 425	32	80.0	170279	13	ABD32686	Abd32686 Mouse can	c 498	31	77.5	471	13	ADV84060	Adv84060 Streptoco
C 426	32	80.0	173810	6	ABN85752	Abn85752 Mouse chr	c 499	31	77.5	476	4	AAAT43458	Aat43458 ATM gene
C 427	32	80.0	177851	8	AAAL57272	Aal57272 bA38B23-	500	31	77.5	476	4	AAAF57610	Aaf57610 ATM genom
C 428	32	80.0	213251	6	ABQ67193	Abq67193 Listeria	501	31	77.5	480	10	ADC75278	Adc75278 T harzian
C 429	32	80.0	249878	10	ACR65381	AcR65381 Photorhab	502	31	77.5	485	12	ACH92171	Ach92171 Human gen
C 430	32	80.0	273254	3	AAH64966	Aah64966 C glutami	c 503	31	77.5	491	12	ADL08466	Adl08466 Human tum
C 431	32	80.0	349980	5	AAH64966	Aah64966 C glutami	c 504	31	77.5	495	6	ABN19447	Abn19447 Human ORF
C 432	31	77.5	94	4	ABSA43292	AbS43292 Human liv	c 505	31	77.5	504	12	ADK16192	Adk16192 Nanoarcha
C 433	31	77.5	102	3	AAAC66501	Aac66501 Chimaeric	506	31	77.5	507	6	ABQ90526	Abq90526 M. capaul
C 434	31	77.5	102	3	AAAC66500	Aac66500 Chimaeric	507	31	77.5	510	3	AZ43765	Aaz43765 Human PCR
C 435	31	77.5	102	3	AAAC68028	Aac68028 Oligonuc	508	31	77.5	510	12	ADO63154	Ado63154 Transcrip
C 436	31	77.5	102	3	AAAC68029	Aac68029 Oligonuc	c 509	31	77.5	517	11	ACL30706	AcL30706 Rice abio
C 437	31	77.5	102	4	AAH24529	Aah24529 Human pri	510	31	77.5	522	4	AAK59685	Aak59685 Human imm
C 438	31	77.5	102	4	AAH24530	Aah24530 Human pri	511	31	77.5	522	8	ACC79989	Acc79989 P. papata
C 439	31	77.5	102	5	AAH24457	Aah24457 Human pri	512	31	77.5	527	6	ABK78855	Abk78855 Bacillus
C 440	31	77.5	102	5	AAH24458	Aah24458 Human pri	513	31	77.5	531	10	ADG37987	Adg37987 Aspergill
C 441	31	77.5	102	5	AAH24874	Aah24874 Oligonuc	514	31	77.5	534	10	ADC75845	Adc75845 DNA homol
C 442	31	77.5	102	5	AAH24875	Aah24875 Oligonuc	515	31	77.5	537	10	ADC77051	Adc77051 DNA homol

c 516	31	77.5	538	9	ACH44608	Ach44608 Human foe	c 589	31	77.5	1106	8	ACD22050	AcD22050 Human sec
c 517	31	77.5	551	12	ACH75760	Ach75760 Human gen	c 590	31	77.5	1106	8	ACF13215	AcF13215 Human sec
c 518	31	77.5	553	12	ACH78467	Ach78467 Human gen	c 591	31	77.5	1106	8	ACD25318	AcD25318 Human sec
c 519	31	77.5	570	6	ABQ90527	Abq90527 M. capsul	c 592	31	77.5	1106	8	ACF00367	AcF00367 Human sec
c 520	31	77.5	575	5	AAS77846	Aas77846 DNA encod	c 593	31	77.5	1106	8	ACF74244	AcF74244 Novel hum
c 521	31	77.5	580	10	ADC75214	Adc75214 T harzian	c 594	31	77.5	1106	8	ACD04948	AcD04948 Novel hum
c 522	31	77.5	581	13	ACN56313	Acn56313 Cotton an	c 595	31	77.5	1106	8	ACD18409	AcD18409 Human sec
c 523	31	77.5	596	6	ABA94947	Abas94947 Soybean K	c 596	31	77.5	1106	8	ACD08416	AcD08416 Human sec
c 524	31	77.5	604	4	AAH71558	Aah71558 Human cer	c 597	31	77.5	1106	8	ACA88850	AcA88850 Novel hum
c 525	31	77.5	647	12	ADQ97850	Adq97850 Mouse can	c 598	31	77.5	1106	8	ACA70292	AcA70292 Human sec
c 526	31	77.5	659	8	ABT33953	Abt33953 Human pig	c 599	31	77.5	1106	8	ACD12514	AcD12514 Novel hum
c 527	31	77.5	660	3	AAZ43763	Aaz43763 Human par	c 600	31	77.5	1106	8	ACC74429	Acc74429 Human sec
c 528	31	77.5	662	13	ADQ48875	Adq48875 Novel can	c 601	31	77.5	1106	8	ACD16057	AcD16057 Human sec
c 529	31	77.5	672	4	AA541230	Aas41230 cDNA enco	c 602	31	77.5	1106	8	ACD25625	AcD25625 Novel hum
c 530	31	77.5	673	3	AAf14243	Aaf14243 Aspergill	c 603	31	77.5	1106	8	ACD18102	AcD18102 Human sec
c 531	31	77.5	673	13	ADU58284	Adu58284 Aspergill	c 604	31	77.5	1106	8	ACC88389	Acc88389 Human sec
c 532	31	77.5	673	14	ADZ96287	Adz96287 Aspergill	c 605	31	77.5	1106	8	ACD21743	AcD21743 Human sec
c 533	31	77.5	675	5	ABV20903	Abv20903 Human pro	c 606	31	77.5	1106	8	ACD18810	AcD18810 Human sec
c 534	31	77.5	682	11	ACL31875	Acl31875 Rice abio	c 607	31	77.5	1106	8	ABX98420	Abx98420 Human sec
c 535	31	77.5	684	11	ACH98583	Ach98583 Klebsiell	c 608	31	77.5	1106	8	ACD14171	AcD14171 Human PRO
c 536	31	77.5	687	8	ACA27211	AcA27211 Prokaryot	c 609	31	77.5	1106	8	ACD09951	AcD09951 Human sec
c 537	31	77.5	696	3	AAc76478	Aac76478 Human ORP	c 610	31	77.5	1106	8	ACC88696	Acc88696 Human sec
c 538	31	77.5	707	6	ABQ32571	Abq32571 Oligonuc	c 611	31	77.5	1106	8	ACD21436	AcD21436 Human sec
c 539	31	77.5	707	6	ABQ32570	Abq32570 Oligonuc	c 612	31	77.5	1106	8	ABX75808	Abx75808 Human sec
c 540	31	77.5	708	9	ADA30933	Ada30933 DNA encod	c 613	31	77.5	1106	8	ABX98011	Abx98011 Human PRO
c 541	31	77.5	709	2	AAQ51037	Aaq51037 Double st	c 614	31	77.5	1106	8	ACA97487	AcA97487 Novel hum
c 542	31	77.5	723	3	AAZ38632	Aaz38632 Human Ig-	c 615	31	77.5	1106	8	ACA57950	AcA57950 Human PRO
c 543	31	77.5	725	6	AD341124	Ad341124 Human sec	c 616	31	77.5	1106	8	ACD14478	AcD14478 Human PRO
c 544	31	77.5	753	10	ADF00030	Adf00030 Bacterial	c 617	31	77.5	1106	8	ACC91261	Acc91261 Human sec
c 545	31	77.5	756	5	ABV27764	Abv27764 Human pro	c 618	31	77.5	1106	8	ACC89003	Acc89003 Human sec
c 546	31	77.5	756	5	ABV21932	Abv21932 Human pro	c 619	31	77.5	1106	8	ACD07200	AcD07200 Human PRO
c 547	31	77.5	783	10	ADH82560	Adh82560 Enterococ	c 620	31	77.5	1106	8	ACA67651	AcA67651 Human PRO
c 548	31	77.5	786	13	ADR62576	Adr62576 Cotton cd	c 621	31	77.5	1106	8	ACC81706	Acc81706 Human sec
c 549	31	77.5	795	2	AAZ15505	Aaz15505 Human gen	c 622	31	77.5	1106	8	ACC89310	Acc89310 Human sec
c 550	31	77.5	801	4	AA515167	Aa515167 Pseudomon	c 623	31	77.5	1106	8	ACC86666	Acc86666 Human sec
c 551	31	77.5	801	4	ACA19413	Aca19413 Prokaryot	c 624	31	77.5	1106	8	ACC89924	Acc89924 Human sec
c 552	31	77.5	811	2	AAV47532	Aav47532 cDNA enco	c 625	31	77.5	1106	8	ACC93103	Acc93103 Human sec
c 553	31	77.5	813	4	AAH03443	Aah03443 Human cDN	c 626	31	77.5	1106	8	ACA72731	AcA72731 Human PRO
c 554	31	77.5	849	8	ACA53909	AcA53909 Prokaryot	c 627	31	77.5	1106	8	ACA89249	AcA89249 Human sec
c 555	31	77.5	906	8	ACA46830	Aca46830 Prokaryot	c 628	31	77.5	1106	8	ACA69985	AcA69985 Human sec
c 556	31	77.5	909	4	AAH52875	Aah52875 S. epider	c 629	31	77.5	1106	8	ACA97128	AcA97128 Novel hum
c 557	31	77.5	918	6	ABN93332	Abn93332 Staphyloc	c 630	31	77.5	1106	8	ACA91124	AcA91124 Novel hum
c 558	31	77.5	918	8	ACA47074	AcA47074 Prokaryot	c 631	31	77.5	1106	8	ACA70906	AcA70906 Human sec
c 559	31	77.5	918	13	ADS04340	Ads04340 Staphyloc	c 632	31	77.5	1106	8	ACA95416	AcA95416 Novel hum
c 560	31	77.5	948	6	ABN91841	Abn91841 Staphyloc	c 633	31	77.5	1106	8	ACC86359	Acc86359 Human sec
c 561	31	77.5	948	13	ADS01097	Ads01097 Staphyloc	c 634	31	77.5	1106	8	ACC90231	Acc90231 Human sec
c 562	31	77.5	951	6	ABK74865	Abk74865 Bacillus	c 635	31	77.5	1106	8	ACD12839	AcD12839 Human sec
c 563	31	77.5	951	8	ACA28951	AcA28951 Prokaryot	c 636	31	77.5	1106	8	ACF20069	AcF20069 Human sec
c 564	31	77.5	954	6	ABK74858	Abk74858 Bacillus	c 637	31	77.5	1106	8	ABX77013	Abx77013 Human PRO
c 565	31	77.5	958	14	ADZ26714	Adz26714 Human CD7	c 638	31	77.5	1106	8	ACA73345	AcA73345 Novel hum
c 566	31	77.5	966	6	ABQ90575	Abq90575 M. capsul	c 639	31	77.5	1106	8	ACA68888	AcA68888 Novel hum
c 567	31	77.5	966	10	ADC26196	Adc26196 Human NOV	c 640	31	77.5	1106	8	ACA74732	AcA74732 cDNA enco
c 568	31	77.5	995	2	AAx14581	Aax14581 H. pylori	c 641	31	77.5	1106	8	ACA70599	AcA70599 Human sec
c 569	31	77.5	999	10	ACF70889	AcF70889 Photorhab	c 642	31	77.5	1106	8	ACD14785	AcD14785 Human PRO
c 570	31	77.5	1005	10	ADC26198	Adc26198 Human NOV	c 643	31	77.5	1106	8	ACA68457	AcA68457 Novel hum
c 571	31	77.5	1026	11	ABD03133	Abd03133 Pseudomon	c 644	31	77.5	1106	8	ABX98922	Abx98922 Novel hum
c 572	31	77.5	1031	12	ADJ41281	Adj41281 Plant cDN	c 645	31	77.5	1106	8	ACC81399	Acc81399 Human sec
c 573	31	77.5	1036	4	AA536104	Aas36104 Human car	c 646	31	77.5	1106	8	ACA95723	AcA95723 Novel hum
c 574	31	77.5	1036	4	AAK83327	Aak83327 Human imm	c 647	31	77.5	1106	8	ACD04641	AcD04641 Novel hum
c 575	31	77.5	1036	10	ADZ46798	Adz46798 Human car	c 648	31	77.5	1106	8	ACC88082	Acc88082 Human sec
c 576	31	77.5	1036	13	ADJ08216	Adj08216 Human car	c 649	31	77.5	1106	8	ACF12744	AcF12744 Human sec
c 577	31	77.5	1038	11	ABD17883	Abd17883 Pseudomon	c 650	31	77.5	1106	8	ACA96459	AcA96459 Human PRO
c 578	31	77.5	1038	11	ABD17512	Abd17512 Pseudomon	c 651	31	77.5	1106	8	ACA65233	AcA65233 Human PRO
c 579	31	77.5	1068	8	ACA18454	AcA18454 Prokaryot	c 652	31	77.5	1106	8	ACA73959	AcA73959 Human sec
c 580	31	77.5	1071	4	AA551238	Aa551238 Enterococ	c 653	31	77.5	1106	8	ACA74371	AcA74371 Novel hum
c 581	31	77.5	1071	4	AA553048	Aa553048 Enterococ	c 654	31	77.5	1106	8	ACA96766	AcA96766 Human PRO
c 582	31	77.5	1106	4	AA546192	Aa546192 Human DNA	c 655	31	77.5	1106	8	ACD10872	AcD10872 cDNA enco
c 583	31	77.5	1106	8	ACA89642	AcA89642 cDNA enco	c 656	31	77.5	1106	8	ACC91568	Acc91568 Human sec
c 584	31	77.5	1106	8	ACA73652	AcA73652 Human sec	c 657	31	77.5	1106	8	ACD02903	AcD02903 cDNA enco
c 585	31	77.5	1106	8	ACA05967	AcA05967 Human sec	c 658	31	77.5	1106	8	ACC87468	Acc87468 Human sec
c 586	31	77.5	1106	8	ACA66801	AcA66801 cDNA enco	c 659	31	77.5	1106	8	ACC86052	Acc86052 Human sec
c 587	31	77.5	1106	8	ACF20376	AcF20376 Human sec	c 660	31	77.5	1106	8	ACA65540	AcA65540 Human PRO
c 588	31	77.5	1106	8	ACF19762	AcF19762 Human sec	c 661	31	77.5	1106	8	ACA94357	AcA94357 Human sec

c 662	31	77.5	1106	8	ACA98101	ACA98101 Human PRO	c 735	31	77.5	1106	9	ACF15364	ACf15364 Human sec
c 663	31	77.5	1106	8	ACA91603	ACA91603 Novel hum	c 736	31	77.5	1106	9	ACC97459	ACC97459 Human sec
c 664	31	77.5	1106	8	ACA90817	ACA90817 Novel hum	c 737	31	77.5	1106	9	ACC92489	ACC92489 Human sec
c 665	31	77.5	1106	8	ACD16364	ACD16364 Human sec	c 738	31	77.5	1106	9	ACF14136	ACf14136 Human sec
c 666	31	77.5	1106	8	ACD17525	ACD17525 Human sec	c 739	31	77.5	1106	9	ACF14443	ACf14443 Human sec
c 667	31	77.5	1106	8	ACD92182	ACD92182 Human sec	c 740	31	77.5	1106	9	ACF09674	ACf09674 Human sec
c 668	31	77.5	1106	8	ACA75039	ACA75039 CDNA enco	c 741	31	77.5	1106	9	ACD45965	ACD45965 Human sec
c 669	31	77.5	1106	8	ACA91910	ACA91910 Human PRO	c 742	31	77.5	1106	9	ACD48114	ACD48114 Human sec
c 670	31	77.5	1106	8	ACA71554	ACA71554 Human sec	c 743	31	77.5	1106	9	ACD67845	ACD67845 CDNA enco
c 671	31	77.5	1106	8	ACA90954	ACA90954 Human sec	c 744	31	77.5	1106	9	ACF25653	ACf25653 Human sec
c 672	31	77.5	1106	8	ACA65964	ACA65964 CDNA enco	c 745	31	77.5	1106	9	ACF29337	ACf29337 Human sec
c 673	31	77.5	1106	8	ACA95109	ACA95109 CDNA enco	c 746	31	77.5	1106	9	ACD85115	ACD85115 Human sec
c 674	31	77.5	1106	8	ACD16671	ACD16671 Human sec	c 747	31	77.5	1106	9	ACD84194	ACD84194 Human PRO
c 675	31	77.5	1106	8	ACD15750	ACD15750 Human sec	c 748	31	77.5	1106	9	ACD88185	ACD88185 Human sec
c 676	31	77.5	1106	8	ABX15853	ABX15853 Human CDN	c 749	31	77.5	1106	9	ACF30872	ACf30872 Human sec
c 677	31	77.5	1106	8	ACA97794	ACA97794 Human PRO	c 750	31	77.5	1106	9	ACF32475	ACf32475 Human sec
c 678	31	77.5	1106	9	ACA99243	ACA99243 Novel hum	c 751	31	77.5	1106	9	ACH12135	ACH12135 CDNA enco
c 679	31	77.5	1106	9	ACC91875	ACC91875 Human sec	c 752	31	77.5	1106	9	ACH12442	ACH12442 CDNA enco
c 680	31	77.5	1106	9	ACD11286	ACD11286 Novel hum	c 753	31	77.5	1106	9	ACD40834	ACD40834 Human sec
c 681	31	77.5	1106	9	ACD15136	ACD15136 Human sec	c 754	31	77.5	1106	9	ACF18306	ACf18306 Human sec
c 682	31	77.5	1106	9	ACD11900	ACD11900 Human sec	c 755	31	77.5	1106	9	ACF08753	ACf08753 Human sec
c 683	31	77.5	1106	9	ACC96029	ACC96029 Human sec	c 756	31	77.5	1106	9	ACF31554	ACf31554 Human sec
c 684	31	77.5	1106	9	ACF16592	ACF16592 Human sec	c 757	31	77.5	1106	9	ACF52394	ACf52394 Human sec
c 685	31	77.5	1106	9	ACF02710	ACF02710 Human sec	c 758	31	77.5	1106	9	ACD50263	ACD50263 Human sec
c 686	31	77.5	1106	9	ACF03017	ACF03017 Human sec	c 759	31	77.5	1106	9	ACF38966	ACf38966 Human sec
c 687	31	77.5	1106	9	ACF21604	ACF21604 Human sec	c 760	31	77.5	1106	9	ACF26881	ACf26881 Human sec
c 688	31	77.5	1106	9	ACF10288	ACF10288 Human sec	c 761	31	77.5	1106	9	ACF24981	ACf24981 Human sec
c 689	31	77.5	1106	9	ACF78181	ACF78181 Human sec	c 762	31	77.5	1106	9	ACF46561	ACf46561 Human sec
c 690	31	77.5	1106	9	ACD46886	ACD46886 Human sec	c 763	31	77.5	1106	9	ACF28109	ACf28109 Human sec
c 691	31	77.5	1106	9	ACD49649	ACD49649 Human sec	c 764	31	77.5	1106	9	ACD89413	ACD89413 Human sec
c 692	31	77.5	1106	9	ACF28416	ACF28416 Human sec	c 765	31	77.5	1106	9	ACF63985	ACf63985 Human sec
c 693	31	77.5	1106	9	ACD89106	ACD89106 Human sec	c 766	31	77.5	1106	9	ACF60625	ACf60625 Human sec
c 694	31	77.5	1106	9	ACD84501	ACD84501 Human PRO	c 767	31	77.5	1106	9	ACH12749	ACH12749 CDNA enco
c 695	31	77.5	1106	9	ACD99275	ACD99275 CDNA enco	c 768	31	77.5	1106	9	ACH10172	ACH10172 Human sec
c 696	31	77.5	1106	9	ADA78287	ADA78287 Human sec	c 769	31	77.5	1106	9	ACD04027	ACD04027 Human sec
c 697	31	77.5	1106	9	ACF49017	ACF49017 Human sec	c 770	31	77.5	1106	9	ACD10565	ACD10565 Human sec
c 698	31	77.5	1106	9	ACD09337	ACD09337 Human sec	c 771	31	77.5	1106	9	ACF42592	ACf42592 Human sec
c 699	31	77.5	1106	9	ACF12130	ACF12130 Human sec	c 772	31	77.5	1106	9	ACF18613	ACf18613 Human sec
c 700	31	77.5	1106	9	ACF41364	ACF41364 Human sec	c 773	31	77.5	1106	9	ACF02403	ACf02403 Human sec
c 701	31	77.5	1106	9	ACF15978	ACF15978 Human sec	c 774	31	77.5	1106	9	ACF21911	ACf21911 Human sec
c 702	31	77.5	1106	9	ACF16285	ACF16285 Human sec	c 775	31	77.5	1106	9	ACF10595	ACf10595 Human sec
c 703	31	77.5	1106	9	ACD32112	ACD32112 Human sec	c 776	31	77.5	1106	9	ACF34047	ACf34047 Human sec
c 704	31	77.5	1106	9	ACF18920	ACF18920 Human sec	c 777	31	77.5	1106	9	ACF45009	ACf45009 Human sec
c 705	31	77.5	1106	9	ACF09367	ACF09367 Human sec	c 778	31	77.5	1106	9	ACD90541	ACD90541 Human sec
c 706	31	77.5	1106	9	ACF78488	ACF78488 Human sec	c 779	31	77.5	1106	9	ACD91254	ACD91254 Human sec
c 707	31	77.5	1106	9	ACF52087	ACF52087 Human sec	c 780	31	77.5	1106	9	ACF30565	ACf30565 Human sec
c 708	31	77.5	1106	9	ACF26574	ACF26574 Human sec	c 781	31	77.5	1106	9	ACD87264	ACD87264 Human sec
c 709	31	77.5	1106	9	ACF24367	ACF24367 Human sec	c 782	31	77.5	1106	9	ACF60318	ACf60318 Human sec
c 710	31	77.5	1106	9	ACF63678	ACF63678 Human sec	c 783	31	77.5	1106	9	ACF46868	ACf46868 Human sec
c 711	31	77.5	1106	9	ACF50552	ACF50552 Human sec	c 784	31	77.5	1106	9	ACF75725	ACf75725 Human sec
c 712	31	77.5	1106	9	ACH08023	ACH08023 Human sec	c 785	31	77.5	1106	9	ADA80079	ADA80079 Human sec
c 713	31	77.5	1106	9	ACF13829	ACF13829 Human sec	c 786	31	77.5	1106	9	ACF17385	ACf17385 Human sec
c 714	31	77.5	1106	9	ACD41755	ACD41755 Human sec	c 787	31	77.5	1106	9	ACF23139	ACf23139 Human sec
c 715	31	77.5	1106	9	ACF32168	ACF32168 Human sec	c 788	31	77.5	1106	9	ACF08139	ACf08139 Human sec
c 716	31	77.5	1106	9	ACF23446	ACF23446 Human sec	c 789	31	77.5	1106	9	ACF08446	ACf08446 Human sec
c 717	31	77.5	1106	9	ACF40136	ACF40136 Human sec	c 790	31	77.5	1106	9	ACF40750	ACf40750 Human sec
c 718	31	77.5	1106	9	ACD45658	ACD45658 Human sec	c 791	31	77.5	1106	9	ACF53929	ACf53929 Human sec
c 719	31	77.5	1106	9	ACF53315	ACF53315 Human sec	c 792	31	77.5	1106	9	ACD47193	ACD47193 Human sec
c 720	31	77.5	1106	9	ACF27495	ACF27495 Human sec	c 793	31	77.5	1106	9	ACF48096	ACf48096 Human sec
c 721	31	77.5	1106	9	ACF45333	ACF45333 Human sec	c 794	31	77.5	1106	9	ACF47482	ACf47482 Human sec
c 722	31	77.5	1106	9	ACF23951	ACF23951 Human sec	c 795	31	77.5	1106	9	ACF46254	ACf46254 Human sec
c 723	31	77.5	1106	9	ACD90027	ACD90027 Human sec	c 796	31	77.5	1106	9	ACD86343	ACD86343 Human sec
c 724	31	77.5	1106	9	ACD84808	ACD84808 Human PRO	c 797	31	77.5	1106	9	ACF52701	ACF52701 Human sec
c 725	31	77.5	1106	9	ACD98968	ACD98968 CDNA enco	c 798	31	77.5	1106	9	ACF53008	ACf53008 Human sec
c 726	31	77.5	1106	9	ACF77260	ACF77260 Human sec	c 799	31	77.5	1106	9	ACF65001	ACf65001 Human sec
c 727	31	77.5	1106	9	ACF76953	ACF76953 Human sec	c 800	31	77.5	1106	9	ACF76646	ACf76646 Human sec
c 728	31	77.5	1106	9	ACF49338	ACF49338 Human sec	c 801	31	77.5	1106	9	ACF61546	ACf61546 Human sec
c 729	31	77.5	1106	9	ACF50245	ACF50245 Human sec	c 802	31	77.5	1106	9	ACF61853	ACf61853 Human sec
c 730	31	77.5	1106	9	ACD09644	ACD09644 Human sec	c 803	31	77.5	1106	9	ACD30884	ACD30884 Human sec
c 731	31	77.5	1106	9	ACD08723	ACD08723 Human sec	c 804	31	77.5	1106	9	ACD31805	ACD31805 Human sec
c 732	31	77.5	1106	9	ACF12437	ACF12437 Human sec	c 805	31	77.5	1106	9	ACD32726	ACD32726 Human sec
c 733	31	77.5	1106	9	ACC94945	ACC94945 Human sec	c 806	31	77.5	1106	9	ACF17692	ACf17692 Human sec
c 734	31	77.5	1106	9	ACD22664	ACD22664 Human sec	c 807	31	77.5	1106	9	ACF07525	ACf07525 Human sec

C 808	31	77.5	1106	9	ACF20683	Human sec
C 809	31	77.5	1106	9	ACF20990	Human sec
C 810	31	77.5	1106	9	ACF21297	Human sec
C 811	31	77.5	1106	9	ACD47807	Human sec
C 812	31	77.5	1106	9	ACF47789	Human sec
C 813	31	77.5	1106	9	ACF53622	Human sec
C 814	31	77.5	1106	9	ACD86957	Human sec
C 815	31	77.5	1106	9	ACH05205	CDNA enco
C 816	31	77.5	1106	9	ACF44702	Human sec
C 817	31	77.5	1106	9	ADA81806	Human sec
C 818	31	77.5	1106	9	ACD22357	Human sec
C 819	31	77.5	1106	9	ACD24704	Human sec
C 820	31	77.5	1106	9	ACD39907	CDNA enco
C 821	31	77.5	1106	9	ACD40214	CDNA enco
C 822	31	77.5	1106	9	ACF13522	Human sec
C 823	31	77.5	1106	9	ACF03324	Human sec
C 824	31	77.5	1106	9	ACF78795	Human sec
C 825	31	77.5	1106	9	ACF11516	Human sec
C 826	31	77.5	1106	9	ACF50859	Human sec
C 827	31	77.5	1106	9	ACF34354	Human sec
C 828	31	77.5	1106	9	ACD46579	Human sec
C 829	31	77.5	1106	9	ACD48421	Human sec
C 830	31	77.5	1106	9	ACF27802	Human sec
C 831	31	77.5	1106	9	ACF24674	Human sec
C 832	31	77.5	1106	9	ACD85729	Human sec
C 833	31	77.5	1106	9	ACD90334	Human sec
C 834	31	77.5	1106	9	ACD83387	Human sec
C 835	31	77.5	1106	9	ACF49324	Human sec
C 836	31	77.5	1106	9	ACH07409	Human sec
C 837	31	77.5	1106	9	ACH07716	Human sec
C 838	31	77.5	1106	9	ACH08330	Human sec
C 839	31	77.5	1106	9	ACH11521	CDNA enco
C 840	31	77.5	1106	9	ACH11828	CDNA enco
C 841	31	77.5	1106	9	ACH10479	Human sec
C 842	31	77.5	1106	9	ACF01482	Human sec
C 843	31	77.5	1106	9	ACF41057	Human sec
C 844	31	77.5	1106	9	ACD24397	Human sec
C 845	31	77.5	1106	9	ACD31498	Human sec
C 846	31	77.5	1106	9	ACF17999	Human sec
C 847	31	77.5	1106	9	ACF32782	Human sec
C 848	31	77.5	1106	9	ACF40443	Human sec
C 849	31	77.5	1106	9	ACF48403	Human sec
C 850	31	77.5	1106	9	ACF38352	Human sec
C 851	31	77.5	1106	9	ACF25288	Human sec
C 852	31	77.5	1106	9	ACF27188	Human sec
C 853	31	77.5	1106	9	ACF29644	Human sec
C 854	31	77.5	1106	9	ACD87878	Human sec
C 855	31	77.5	1106	9	ACF76339	Human sec
C 856	31	77.5	1106	9	ACF49631	Human sec
C 857	31	77.5	1106	9	ACF44088	Human sec
C 858	31	77.5	1106	9	ACH06433	CDNA enco
C 859	31	77.5	1106	9	ACH06740	CDNA enco
C 860	31	77.5	1106	9	ADA83604	Human sec
C 861	31	77.5	1106	9	ACC92796	Human sec
C 862	31	77.5	1106	9	ACC93410	Human sec
C 863	31	77.5	1106	9	ACF19455	Human sec
C 864	31	77.5	1106	9	ACD13146	Human sec
C 865	31	77.5	1106	9	ACF06604	Human sec
C 866	31	77.5	1106	9	ACC94638	Human sec
C 867	31	77.5	1106	9	ACC98066	Human sec
C 868	31	77.5	1106	9	ACC94331	Human sec
C 869	31	77.5	1106	9	ACF42285	Human sec
C 870	31	77.5	1106	9	ACD31191	Human sec
C 871	31	77.5	1106	9	ACD43220	CDNA enco
C 872	31	77.5	1106	9	ACD43527	CDNA enco
C 873	31	77.5	1106	9	ACF15057	Human sec
C 874	31	77.5	1106	9	ACF01789	Human sec
C 875	31	77.5	1106	9	ACF31861	Human sec
C 876	31	77.5	1106	9	ACD67538	CDNA enco
C 877	31	77.5	1106	9	ACD48728	Human sec
C 878	31	77.5	1106	9	ACD49035	Human sec
C 879	31	77.5	1106	9	ACF51473	Human sec
C 880	31	77.5	1106	9	ACF54236	Human sec
C 881	31	77.5	1106	9	ACF20683	Human sec
C 882	31	77.5	1106	9	ACF20990	Human sec
C 883	31	77.5	1106	9	ACF21297	Human sec
C 884	31	77.5	1106	9	ACD47807	Human sec
C 885	31	77.5	1106	9	ACF47789	Human sec
C 886	31	77.5	1106	9	ACF53622	Human sec
C 887	31	77.5	1106	9	ACD86957	Human sec
C 888	31	77.5	1106	9	ACH05205	CDNA enco
C 889	31	77.5	1106	9	ACF44702	Human sec
C 890	31	77.5	1106	9	ADA81806	Human sec
C 891	31	77.5	1106	9	ACD22357	Human sec
C 892	31	77.5	1106	9	ACD24704	Human sec
C 893	31	77.5	1106	9	ACD39907	CDNA enco
C 894	31	77.5	1106	9	ACD40214	CDNA enco
C 895	31	77.5	1106	9	ACF13522	Human sec
C 896	31	77.5	1106	9	ACF03324	Human sec
C 897	31	77.5	1106	9	ACF78795	Human sec
C 898	31	77.5	1106	9	ACF11516	Human sec
C 899	31	77.5	1106	9	ACF50859	Human sec
C 900	31	77.5	1106	9	ACF34354	Human sec
C 901	31	77.5	1106	9	ACD46579	Human sec
C 902	31	77.5	1106	9	ACD48421	Human sec
C 903	31	77.5	1106	9	ACF27802	Human sec
C 904	31	77.5	1106	9	ACF24674	Human sec
C 905	31	77.5	1106	9	ACD85729	Human sec
C 906	31	77.5	1106	9	ACD90334	Human sec
C 907	31	77.5	1106	9	ACD83387	Human sec
C 908	31	77.5	1106	9	ACF49324	Human sec
C 909	31	77.5	1106	9	ACH07409	Human sec
C 910	31	77.5	1106	9	ACH07716	Human sec
C 911	31	77.5	1106	9	ACH08330	Human sec
C 912	31	77.5	1106	9	ACH11521	CDNA enco
C 913	31	77.5	1106	9	ACH11828	CDNA enco
C 914	31	77.5	1106	9	ACH10479	Human sec
C 915	31	77.5	1106	9	ACF01482	Human sec
C 916	31	77.5	1106	9	ACF41057	Human sec
C 917	31	77.5	1106	9	ACD24397	Human sec
C 918	31	77.5	1106	9	ACD31498	Human sec
C 919	31	77.5	1106	9	ACF17999	Human sec
C 920	31	77.5	1106	9	ACF32782	Human sec
C 921	31	77.5	1106	9	ACF40443	Human sec
C 922	31	77.5	1106	9	ACF48403	Human sec
C 923	31	77.5	1106	9	ACF38352	Human sec
C 924	31	77.5	1106	9	ACF25288	Human sec
C 925	31	77.5	1106	9	ACF27188	Human sec
C 926	31	77.5	1106	9	ACF29644	Human sec
C 927	31	77.5	1106	9	ACD87878	Human sec
C 928	31	77.5	1106	9	ACF76339	Human sec
C 929	31	77.5	1106	9	ACF49631	Human sec
C 930	31	77.5	1106	9	ACF44088	Human sec
C 931	31	77.5	1106	9	ACH06433	CDNA enco
C 932	31	77.5	1106	9	ACH06740	CDNA enco
C 933	31	77.5	1106	9	ADA83604	Human sec
C 934	31	77.5	1106	9	ACC92796	Human sec
C 935	31	77.5	1106	9	ACC93410	Human sec
C 936	31	77.5	1106	9	ACF19455	Human sec
C 937	31	77.5	1106	9	ACD13146	Human sec
C 938	31	77.5	1106	9	ACF06604	Human sec
C 939	31	77.5	1106	9	ACC94638	Human sec
C 940	31	77.5	1106	9	ACC98066	Human sec
C 941	31	77.5	1106	9	ACC94331	Human sec
C 942	31	77.5	1106	9	ACF42285	Human sec
C 943	31	77.5	1106	9	ACD31191	Human sec
C 944	31	77.5	1106	9	ACD43220	CDNA enco
C 945	31	77.5	1106	9	ACD43527	CDNA enco
C 946	31	77.5	1106	9	ACF15057	Human sec
C 947	31	77.5	1106	9	ACF01789	Human sec
C 948	31	77.5	1106	9	ACF31861	Human sec
C 949	31	77.5	1106	9	ACD67538	CDNA enco
C 950	31	77.5	1106	9	ACD48728	Human sec
C 951	31	77.5	1106	9	ACD49035	Human sec
C 952	31	77.5	1106	9	ACF51473	Human sec
C 953	31	77.5	1106	9	ACF54236	Human sec

XX The invention describes a method comprising comparing an expression
 CC profile of at least one gene in a peripheral blood sample of a patient to
 CC at least one reference expression profile of the at least one gene, where
 CC the patient has a solid tumour, and each of the gene is differentially
 CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
 CC of patients as compared to PBMCs of a second class of patients, where
 CC both the first and second classes of patients have the solid tumour, and
 CC each of the first and second classes is a subcluster formed by an
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of
 CC a population of patients who have the solid tumour, and where the
 CC majority of the first class of patients has a first clinical outcome, and
 CC the majority of the second class of patients has a second clinical
 CC outcome. Also described are: a system comprising (i) a memory or a
 CC storage medium including data that represent an expression profile of at
 CC least one gene in a peripheral blood sample of a patient who has a solid
 CC tumour, (ii) at least another storage medium including data that
 CC represent at least one reference expression profile of the gene, (iii) a
 CC program capable of comparing the expression profile to the reference
 CC expression profile, and (iv) a processor capable of executing the
 CC program, where expression levels of the gene in peripheral blood
 CC mononuclear cells of patients who have the solid tumour correlate with
 CC clinical outcomes of the patients; and a nucleic acid or protein array
 CC comprising concentrated probes for solid tumour prognosis genes, where
 CC each of the solid tumour prognosis genes is differentially expressed in
 CC PBMCs of a first class of patients as compared to PBMCs of a second class
 CC of patients, where both the first and second classes of patients have a
 CC solid tumour, and where the first class of patients has a first clinical
 CC outcome, and the second class of patients has a second clinical outcome.
 CC The method, system, and array are useful for prognosing and treating
 CC solid tumors. This sequence represents a solid tumour prognosis gene of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 31.6 Length: 475
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0

US-10-774-176-6 (1-9) x ABU11677 (1-475)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
 DB 369 GCCCTGATAGCGCTATTCTCTCTG 395

RESULT 3
 AAA27060
 ID AAA27060 standard; DNA; 901 BP.

XX AC AAA27060;

XX DT 22-AUG-2000 (first entry)

XX DE Canine 5T4 tumour-associated antigen gene.

XX Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
 KW ds.

XX OS Canis sp.

XX FH Key Location/Qualifiers

XX CDS 1..858

FT /tag= a

FT /product= "5T4 antigen"

FT misc_feature 61..74

FT /tag= b

FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 135..146
 FT /tag= c
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 207..216
 FT /tag= d
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 277..290
 FT /tag= e
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 351..361
 FT /tag= f
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 422..436
 FT /tag= g
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 497..511
 FT /tag= h
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 572..583
 FT /tag= i
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 644..653
 FT /tag= j
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 714..723
 FT /tag= k
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 784..801
 FT /tag= l
 FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"

WO200029428-A2.

25-MAY-2000.

18-NOV-1999; 99WO-GB003859.

18-NOV-1998; 98GB-00025303.

27-JAN-1999; 99GB-00001739.

30-JUL-1999; 99GB-00017995.

(OXFO-) OXFORD BIOMEDICA UK LTD.

Carroll MW, Myers KA;

WPI; 2000-387735/33.

P-PSDB; AAY94351.

Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 response useful in vaccinating against and in treating tumors.

PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 XX Claim 22; Page 336; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 67.5 Length: 927
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x ABX76333 (1-927)

Qy 1 AlaLeulleGlyAlaIlePheLeuleu 9
 Db 748 GCCCTGATAGGCGCTATTTCCTCTG 774

RESULT 6

ID ADB80503 standard; DNA; 927 BP.

XX ADB80503;

XX 04-DEC-2003 (first entry)

XX Ovarian cancer-associated transcript #34.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; db; gene.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..927

FT /*tag= a

XX WO2002102235-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

XX

PA (EO5B-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

DR P-PSDB; ADB80504.

XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

XX Claim 10; Page 297; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.

XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 67.5 Length: 927
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-6 (1-9) x ADB80503 (1-927)

Qy 1 AlaLeulleGlyAlaIlePheLeuleu 9

Db 748 GCCCTGATAGGCGCTATTTCCTCTG 774

RESULT 7

ADN38723

XX ADN38723 standard; cDNA; 927 BP.

XX ADN38723;

XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

XX WO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 23-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (BOSB-) BOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
DR P-PSDB; ADN38724.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX Claim 8; SEQ ID NO 41; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularization syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 67.5 Length: 927
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-6 (1-9) x ADN38723 (1-927)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9
|||||
Db 748 GCCCTGATAGGGCTATTTCCTCTG 774

RESULT 8
AAD56198
ID AAD56198 standard; DNA; 973 BP.
XX
XX AAD56198;
XX
XX 07-AUG-2003 (first entry)
XX
XX Human LRRCAPS related DNA #5.

XX Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
XX
XX Homo sapiens.
XX WO2003035831-A2.
XX 01-MAY-2003.
XX
XX 21-OCT-2002; 2002WO-US033540.
XX
XX 22-OCT-2001; 2001US-0338733P.
XX 15-FEB-2002; 2002US-0357600P.
XX 01-MAR-2002; 2002US-0361196P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
XX WPI; 2003-421410/39.
XX
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
XX Example 5; Page 74-75; 99pp; English.
XX
XX The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
XX Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 71.4 Length: 973
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x AAD56198 (1-973)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9
|||||
Db 763 GCCCTGATAGGGCTATTTCCTCTG 789

RESULT 9
ABV99349
ID ABV99349 standard; DNA; 1156 BP.
XX
XX ABV99349;
XX
XX 27-JAN-2003 (first entry)
XX
XX Human NOV8a coding sequence.

XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;

KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX Homo sapiens.
XX WO200272771-A2.
XX 19-SEP-2002.
XX 08-MAR-2002; 2002WO-US007288.
XX 08-MAR-2001; 2001US-0274101P.
XX 08-MAR-2001; 2001US-0274194P.
XX 08-MAR-2001; 2001US-0274281P.
XX 08-MAR-2001; 2001US-0274322P.
XX 09-MAR-2001; 2001US-0274849P.
XX 12-MAR-2001; 2001US-0275235P.
XX 13-MAR-2001; 2001US-0275578P.
XX 13-MAR-2001; 2001US-0275579P.
XX 13-MAR-2001; 2001US-0275601P.
XX 14-MAR-2001; 2001US-0276000P.
XX 16-MAR-2001; 2001US-0276776P.
XX 19-MAR-2001; 2001US-0276994P.
XX 20-MAR-2001; 2001US-0277239P.
XX 20-MAR-2001; 2001US-0277321P.
XX 20-MAR-2001; 2001US-0277327P.
XX 20-MAR-2001; 2001US-0277338P.
XX 21-MAR-2001; 2001US-0277791P.
XX 22-MAR-2001; 2001US-0277833P.
XX 23-MAR-2001; 2001US-0278152P.
XX 26-MAR-2001; 2001US-0278894P.
XX 27-MAR-2001; 2001US-0278999P.
XX 27-MAR-2001; 2001US-0279036P.
XX 28-MAR-2001; 2001US-0279344P.
XX 30-MAR-2001; 2001US-0279995P.
XX 30-MAR-2001; 2001US-0280233P.
XX 02-APR-2001; 2001US-0280802P.
XX 02-APR-2001; 2001US-0280822P.
XX 02-APR-2001; 2001US-0280900P.
XX 04-APR-2001; 2001US-0281194P.
XX 13-APR-2001; 2001US-0283675P.
XX 30-APR-2001; 2001US-0287424P.
XX 02-MAY-2001; 2001US-0288066P.
XX 03-MAY-2001; 2001US-0288342P.
XX 15-MAY-2001; 2001US-0291190P.
XX 16-MAY-2001; 2001US-0291099P.
XX 30-MAY-2001; 2001US-0291240P.
XX 31-MAY-2001; 2001US-0294485P.
XX 31-MAY-2001; 2001US-0294889P.
XX 18-JUN-2001; 2001US-0299027P.
XX 19-JUN-2001; 2001US-0299303P.
XX 19-JUN-2001; 2001US-0299310P.
XX 30-JUL-2001; 2001US-0304354P.
XX 31-JUL-2001; 2001US-0309198P.
XX 16-AUG-2001; 2001US-0312903P.
XX 10-SEP-2001; 2001US-0318462P.
XX 12-SEP-2001; 2001US-0318770P.
XX 27-SEP-2001; 2001US-0325430P.
XX 27-SEP-2001; 2001US-0325681P.
XX 18-OCT-2001; 2001US-0330380P.
XX 31-OCT-2001; 2001US-0335301P.
XX 14-NOV-2001; 2001US-0332172P.
XX 14-NOV-2001; 2001US-0332271P.
XX 14-NOV-2001; 2001US-0332272P.
XX 14-NOV-2001; 2001US-0333184P.
XX 14-NOV-2001; 2001US-0333272P.
XX 21-NOV-2001; 2001US-0332094P.
XX 03-DEC-2001; 2001US-0337426P.
XX 03-DEC-2001; 2001US-0338092P.
XX 04-DEC-2001; 2001US-0337185P.

PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX (CURA-) CURAGEN CORP.
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
PI Voss EZ, Malyankar UM, Anderson DW, Paturajan M, Miller CE;
PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VI, Pochart PF;
PI Zhong M;
XX WPI; 2002-732824/79.
DR P-PSDB; ABP70071.
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
PT disorders, and asthma.
XX Claim 16; Page 114-115; 619pp; English.
XX The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, haematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 86.8 Length: 1156
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x ABV99349 (1-1156)
QY 1 AlaLeuileGlyAlaIlePheLeuLeu 9
Db 979 GCCTGATAGGCGCTATTTTCCTCTG 1005

RESULT 10
ABK87175
ID ABK87175 standard; cDNA; 1260 BP.
XX AC ABK87175;
XX 07-OCT-2002 (first entry)
DT cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
XX Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX

```
OS Felis sp.
XX
XX Key Location/Qualifiers
XX CDS 1..1260
XX /*tag= a
XX /product= "574 protein"
XX
XX WO200238612-A2.
XX
XX 16-MAY-2002.
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Myers K, Drury N, Carroll M;
XX
XX WPI; 2002-557449/59.
XX P-PSDB; AAU98694.
XX
XX Novel canine or feline 574 polypeptide and polynucleotides encoding the
XX polypeptide, useful in preparation of vaccine for treating and/or
XX preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 4; Page 68; 68pp; English.
XX
XX The present invention relates to the isolation of canine and feline
XX oncofoetal leucine-rich glycoproteins known as 574, and the
XX polynucleotide sequences encoding them. The 574 proteins are expressed in
XX a significant proportion of tumours. The sequences of the invention are
XX useful in a pharmaceutical composition for the prevention and/or
XX treatment of tumours or other diseases associated with cell
XX proliferation, infections, and inflammatory conditions in animals,
XX preferably dogs or cats. The compositions may also be used for cancer
XX immunotherapy in these animals. The sequences of the invention may also
XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
XX measurement and localisation of 574 in extracts of plasma, urine,
XX tissues, and in cell culture media. Antibodies specific for the 574
XX protein are useful for isolating foetal cells from maternal blood. The
XX isolation process may form part of a diagnostic method e.g. the foetal
XX cells may then be subject to biochemical or genetic sampling used for
XX testing foetal abnormalities, or to determine the sex of the foetus(es).
XX
XX The present sequence encodes feline 574 protein
XX
XX Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 95.7 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x ABK87175 (1-1260)
Qy 1 AlaLeulleGlyAlaallePheLeuLeu 9
Db 1087 GCCCTGATAGTGCCATTTCCTACTG 1113
RESULT 11
ADB97513
ID ADB97513 standard; DNA; 1260 BP.
XX
XX ADB97513;
XX
XX 04-DEC-2003 (first entry)
XX
XX Feline 574 antigen DNA.
XX
XX Major Histocompatibility Complex class I peptide epitope; MHC;

KW 574 antigen; 574 epitope; polypeptide string; vaccine; T cell;
KW cytostatic; cancer; feline; gene; ds.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
XX CDS 1..1260
XX /*tag= a
XX /product= "Feline 574 antigen protein"
XX
XX WO2003068816-A1.
XX
XX 21-AUG-2003.
XX
XX 13-FEB-2003; 2003WO-GB000670.
XX
XX 13-FEB-2002; 2002GB-00003419.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll M, Kingman S, Redchenko I;
XX
XX WPI; 2003-637141/60.
XX P-PSDB; ADB97520.
XX
XX New major histocompatibility complex class I peptide epitopes from human
XX 574 tumor-associated antigen, useful for preventing and/or treating a
XX disease, particularly cancer.
XX
XX Disclosure; Page 67; 73pp; English.
XX
XX The invention relates to a novel Major Histocompatibility Complex (MHC)
XX class I peptide epitope of the 574 antigen. The invention further
XX provides a polypeptide string comprising the 574 epitope; a nucleic acid
XX sequence encoding the 574 epitope or a polypeptide string of the 574
XX epitope; a vector system capable of delivering the 574 epitope nucleic
XX acid to a cell; a cell pulsed with the 574 epitope, a polypeptide of the
XX 574 epitope, its encoding nucleic acid, or the vector system; a vaccine
XX comprising the above; a method for treating and/or preventing a disease
XX in a subject by administering the vaccine; an agent capable of binding
XX specifically to the 574 epitope and/its encoding nucleic acid; a method
XX comprising detecting the presence of the 574 epitope or its encoding
XX nucleic acid in a subject; and a T cell line or clone capable of
XX specifically recognising the 574 epitope in conjunction with an MHC class
XX I molecule. The 574 epitope has cytostatic activity. The vaccine
XX comprising the 574 epitope or its encoding nucleic acid and the vector
XX system or cell is useful in the prevention and/or treatment of a disease,
XX particularly cancer. The detection method is useful for diagnosing or
XX monitoring the progression of a cancerous disease, and for detecting the
XX presence of the 574 epitope or its nucleic acid. The T cell line or clone
XX is useful in the manufacture of a medicament for treating and/or
XX preventing a disease. This polynucleotide sequence represents the feline
XX 574 antigen coding DNA of the invention.
XX
XX Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 95.7 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-6 (1-9) x ADB97513 (1-1260)
Qy 1 AlaLeulleGlyAlaallePheLeuLeu 9
Db 1087 GCCCTGATAGTGCCATTTCCTACTG 1113
RESULT 12
ADB97452
ID ADB97452 standard; DNA; 1260 BP.
```

```

XX AC ADB97452;
XX DT 04-DEC-2003 (first entry)
XX DE DNA encoding feline 5T4 protein.
XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
XX KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
XX OS Unidentified.
XX FH Key Location/Qualifiers
XX FT 1. 1260
XX FT /tag= a
XX FT /product= "feline 5T4 antigen protein"
XX FN WO2003068815-A2.
XX PD 21-AUG-2003.
XX PF 13-FEB-2003; 2003WO-GB000618.
XX PR 13-FEB-2002; 2002GB-00003420.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll M, Harrop R, Kingsman S;
XX WPI; 2003-663795/62.
XX DR P-PSDB; ADB97455.
XX PT New Major Histocompatibility Complex class II peptide epitope of 5T4,
XX FT useful for manufacturing a medicament for diagnosing, preventing and/or
XX FT treating a disease, e.g. cancer.
XX PS Disclosure; Page 49; 63pp; English.
XX CC The invention relates to a Major Histocompatibility Complex (MHC) class
XX CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
XX CC clone has a cytostatic activity, as it is useful in manufacturing a
XX CC medicament for preventing and/or treating a disease, particularly cancer.
XX CC The methods are useful for detecting T-cells capable of specifically
XX CC recognising a peptide epitope in conjunction with an MHC molecule, for
XX CC diagnosing or monitoring the progression of a cancerous disease, or for
XX CC detecting the presence of a peptide or nucleic acid using an agent. The
XX CC MHC class II peptide epitope of the invention can be used in gene therapy
XX CC or as part of a vaccine. This polynucleotide sequence represents the DNA
XX CC coding for the feline 5T4 protein.
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 95.7 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-6 (1-9) x ADB97452 (1-1260)
OY 1 AlaLeuileGlyAlailePheLeuLeu 9
DB 1087 GCCCTGATAGGCGCCATTTCCTACTG 1113

RESULT 13
AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX AC AAA27058;
XX DT 22-AUG-2000 (first entry)

```

```

XX DE Human 5T4 tumour-associated antigen gene.
XX KW Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX KW ds.
XX OS Homo sapiens.
XX PN WO2000029428-A2.
XX PD 25-MAY-2000.
XX PF 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MW, Myers KA;
XX WPI; 2000-387735/33.
XX PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX PS response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the human 5T4 tumour-associated antigen
XX CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The 5T4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject
XX SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 96 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-6 (1-9) x AAA27058 (1-1263)
OY 1 AlaLeuileGlyAlailePheLeuLeu 9
DB 1090 GCCCTGATAGGCGCTATTTCCTCTCG 1116

RESULT 14
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX AC AAF89736;
XX DT 23-JUL-2001 (first entry)
XX DE Nucleotide sequence of canine 5T4 protein.
XX KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
XX KW hypersensitivity; autoimmune disease; central nervous system disorder;
XX KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
XX KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
XX KW Helicobacter-related disease; immune disorder; ss.

```

XX OS Canis sp.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..1263
 XX FT /*tag= a
 XX FT /product= "5T4"
 XX PN WO200136486-A2.
 XX PD 25-MAY-2001.
 XX PF 13-NOV-2000; 2000WO-GB004317.
 XX PR 18-NOV-1999; 99WO-GB003859.
 XX PR 15-FEB-2000; 2000GB-00003527.
 XX PR 02-MAR-2000; 2000GB-00005071.
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PI Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard PM;
 XX PI Myers KA;
 XX DR WPI; 2001-343805/36.
 XX DR P-PSDB; AAB83839.
 XX XX
 XX PT Use of single chain antibody capable of recognizing a disease associated
 XX PT molecule for manufacturing a medicament for preventing and/or treating a
 XX PT disease condition associated with disease associated molecule.
 XX XX
 XX PS Disclosure; Fig 26; 118pp; English.
 XX CC The specification describes the use of a single chain antibody (ScFv),
 XX CC which is capable of recognizing a disease associated molecule in the
 XX CC manufacture of a medicament for the prevention and treatment of a disease
 XX CC condition. The ScFv antibody is useful in the manufacture of a
 XX CC medicament, for affecting a disease in vivo, for preparing a
 XX CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
 XX CC treatment of a disease. The ScFv antibody is also useful for treating
 XX CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
 XX CC diseases, cancers, central nervous system disorders including Parkinson's
 XX CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
 XX CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
 XX CC related diseases, and other immune disorders. The present sequence
 XX CC encodes a 5T4 protein, which is used to produce ScFv of the invention
 XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 96 Length: 1263
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-6 (1-9) x AAF89736 (1-1263)
 Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
 Db 1090 GCCCTGATAGGCCCATCTCTACTG 1116
 RESULT 15
 ABK87174
 ID ABK87174 standard; cDNA; 1263 BP.
 XX AC ABK87174;
 XX XX
 XX DT 07-OCT-2002 (first entry)
 XX DE cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
 XX XX
 XX KW Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;

KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX OS Canis sp.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..1263
 XX FT /*tag= a
 XX FT /product= "5T4 protein"
 XX PN WO200238612-A2.
 XX PD 16-MAY-2002.
 XX PF 13-NOV-2001; 2001WO-GB005004.
 XX PR 13-NOV-2000; 2000WO-GB004317.
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PI Myers K, Drury N, Carroll M;
 XX PI WPI; 2002-557449/59.
 XX DR P-PSDB; AAU98693.
 XX XX
 XX PT Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 XX PT polypeptide, useful in preparation of vaccine for treating and/or
 XX PT preventing cancer in a subject, preferably a dog or cat.
 XX XX
 XX PS Claim 1; Page 67; 68pp; English.
 XX CC The present invention relates to the isolation of canine and feline
 XX CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 XX CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 XX CC a significant proportion of tumours. The sequences of the invention are
 XX CC useful in a pharmaceutical composition for the prevention and/or
 XX CC treatment of tumours or other diseases associated with cell
 XX CC proliferation, infections, and inflammatory conditions in animals,
 XX CC preferably dogs or cats. The compositions may also be used for cancer
 XX CC immunotherapy in these animals. The sequences of the invention may also
 XX CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 XX CC measurement and localisation of 5T4 in extracts of plasma, urine,
 XX CC tissues, and in cell culture media. Antibodies specific for the 5T4
 XX CC protein are useful for isolating foetal cells from maternal blood. The
 XX CC isolation process may form part of a diagnostic method e.g. the foetal
 XX CC cells may then be subject to biochemical or genetic sampling used for
 XX CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 XX CC The present sequence encodes canine 5T4 protein
 XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 96 Length: 1263
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-6 (1-9) x ABK87174 (1-1263)
 Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
 Db 1090 GCCCTGATAGGCCCATCTCTACTG 1116
 RESULT 16
 AAA27059
 ID AAA27059 standard; DNA; 1281 BP.
 XX AC AAA27059;
 XX AC AAA27059;
 XX XX
 XX DT 22-AUG-2000 (first entry)

```

XX DE Mouse ST4 tumour-associated antigen gene.
XX KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX OS Mus musculus.
XX PN WO200029428-A2.
XX PD 25-MAY-2000.
XX PF 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MW, Myers KA;
XX WPI; 2000-387735/33.
XX DR Tumor associated antigen, ST4 capable of eliciting cytotoxic T-lymphocyte
XX PT response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the mouse ST4 tumour-associated antigen
XX CC (TAA). The TAA ST4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. ST4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The ST4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject. The
XX CC present sequence appears in GenBank at accession number AJ012160
XX SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 97.6 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-6 (1-9) x AAA27059 (1-1281)
QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1108 GCTCTGATAGGCGCTATTTCCTCTC 1134

RESULT 17
AAD56199
ID AAD56199 standard; DNA; 1331 BP.
XX AC AAD56199;
XX DT 07-AUG-2003 (first entry)
XX DE Human LRRCAPS related DNA #6.
XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
XX KW LRRCAPS; cancer; gene therapy; ds.
XX OS Homo sapiens.

XX DE Mouse ST4 tumour-associated antigen gene.
XX KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX OS Mus musculus.
XX PN WO200029428-A2.
XX PD 25-MAY-2000.
XX PF 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MW, Myers KA;
XX WPI; 2000-387735/33.
XX DR Tumor associated antigen, ST4 capable of eliciting cytotoxic T-lymphocyte
XX PT response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the mouse ST4 tumour-associated antigen
XX CC (TAA). The TAA ST4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. ST4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The ST4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject. The
XX CC present sequence appears in GenBank at accession number AJ012160
XX SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 97.6 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-6 (1-9) x AAA27059 (1-1281)
QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1108 GCTCTGATAGGCGCTATTTCCTCTC 1134

RESULT 17
AAD56199
ID AAD56199 standard; DNA; 1331 BP.
XX AC AAD56199;
XX DT 07-AUG-2003 (first entry)
XX DE Human LRRCAPS related DNA #6.
XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
XX KW LRRCAPS; cancer; gene therapy; ds.
XX OS Homo sapiens.

```

```

XX WO2003035831-A2.
XX 01-MAY-2003.
XX PF 21-OCT-2002; 2002WO-US033540.
XX PR 22-OCT-2001; 2001US-0338733P.
XX PR 15-FEB-2002; 2002US-0357600P.
XX PR 01-MAR-2002; 2002US-0361196P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX WPI; 2003-421410/39.
XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
XX PT comprises contacting an assay system comprising a purified leucine rich
XX PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX PS Disclosure; Page 75-76; 99pp; English.
XX CC The invention relates to a method of identifying a candidate p53 pathway
XX CC modulating agent. The method involves contacting an assay system
XX CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX CC polypeptide or nucleic acid or its fragment with a test agent and
XX CC detecting a test agent-biased activity, where a difference between the
XX CC test agent-biased activity and the reference activity identifies the test
XX CC agent as a candidate p53 pathway modulating agent. The method is useful
XX CC for identifying a candidate p53 pathway-modulating agent for preparing a
XX CC composition for diagnosing or treating cancer. The invention is useful in
XX CC gene therapy. The present sequence is human LRRCAPS related DNA
XX SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 102 Length: 1331
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x AAD56199 (1-1331)
QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1120 GCCCTGATAGGCGCTATTTCCTCTCG 1146

RESULT 18
ADJ56299
ID ADJ56299 standard; cDNA; 2020 BP.
XX AC ADJ56299;
XX DT 06-MAY-2004 (first entry)
XX DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.
XX KW human; differential expression; transactivator; proto-oncogene;
XX KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
XX KW MYCN activated cell.
XX OS Homo sapiens.
XX PN US2003119009-A1.
XX PD 26-JUN-2003.
XX PR 25-FEB-2002; 2002US-00084817.
XX OS

```


PR 23-FEB-2001; 2001US-0270784P.
 XX (STUA/) STUART S G.
 PA (NUCH/) NUCHTERN J G.
 PA (PLON/) PLON S E.
 PA (SHOH/) SHOHET J M.
 XX
 XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
 PI WPI; 2003-635698/60.
 XX
 XX New genes regulated by MYCN activation, useful in gene therapy,
 PT particularly for treating a subject with e.g. neuroblastoma or other
 PT cancers, or for diagnosing, staging or monitoring the treatment of the
 PT cancer.
 XX
 XX Claim 1; SEQ ID NO 105; 27pp; English.
 XX
 XX This invention relates to novel isolated cDNAs that are differentially
 CC expressed in MYCN activated cells. Specifically, it refers to
 CC polynucleotide sequences that exhibit differential expression patterns in
 CC cells activated by the transactivator MYCN, where MYCN is a proto-
 CC oncogene that is amplified in neuroblastoma cells and is common in small
 CC cell lung cancers. The present invention describes these cDNA molecules
 CC as useful for in hybridisation assays to detect expression of nucleic
 CC acids (or complementary nucleic acids) in a present in a given sample, as
 CC well as for screening assays by identifying molecules or compounds that
 CC specifically bind the cDNA as a ligand and modulate function or activity.
 CC Accordingly, these compositions exhibit cytostatic activity and can also
 CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
 CC that is differentially expressed in MYCN activated cells, given in an
 CC exemplification of the invention. NOTE: This sequence does not appear in
 CC the printed specification but has been obtained in electronic format from
 CC the US Patent Office at
 CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
 XX
 XX Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 164 Length: 2020
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-6 (1-9) x ADJ56299 (1-2020)
 Qy 1 AlaleuileGlyAlaIlePheLeuLeu 9
 Db 1160 GCCCTGATAGCGCTATTTCCTCTG 1186
 RESULT 19
 ACC51052
 ID ACC51052 standard; cDNA; 2053 BP.
 XX
 XX ACC51052;
 XX
 XX 12-JUN-2003 (first entry)
 DT
 XX
 XX Human bladder cancer associated cDNA sequence SEQ ID NO:192.
 DE
 XX Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO2003003906-A2.
 PN
 XX 16-JAN-2003.
 PD
 XX 03-JUL-2002; 2002WO-US021338.
 PF
 XX 03-JUL-2001; 2001US-0302814P.
 PR

PR 03-AUG-2001; 2001US-0310099P.
 PR 08-NOV-2001; 2001US-0343705P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Mack DH, Aziz N;
 PI WPI; 2003-201532/19.
 DR P-FSDB; ABR48236.
 DR
 XX Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT bladder cancer-associated polynucleotide or antibody.
 XX
 XX Claim 6; Page 296; 307pp; English.
 XX
 XX The present invention describes a method for detecting a bladder cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridises to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications
 XX
 XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 167 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-6 (1-9) x ACC51052 (1-2053)
 Qy 1 AlaleuileGlyAlaIlePheLeuLeu 9
 Db 1174 GCCCTGATAGCGCTATTTCCTCTG 1200
 RESULT 20
 ABX76332
 ID ABX76332 standard; DNA; 2053 BP.
 XX
 XX ABX76332;
 XX
 XX 02-APR-2003 (first entry)
 DT
 XX
 XX Lung cancer-associated polynucleotide #196.
 DE
 XX
 XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 XX Unidentified.
 OS
 XX WO200286443-A2.
 PN
 XX 31-OCT-2002.
 PD
 XX 18-APR-2002; 2002WO-US012476.
 PF

XX 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N, Murray R;
 XX
 DR WPI; 2003-093161/08.
 DR P-PSDB; ABUS6603.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 PS Claim 22; Page 335; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 167 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-6 (1-9) x ABX76332 (1-2053)
 Qy 1 AlaLeuileGlyAlailePheLeuLeu 9
 Db 1174 GCCCTGATAGGCGCTATTTCCTCTCG 1200
 RESULT 21
 AAD56197
 ID AAD56197 standard; DNA; 2053 BP.
 XX
 AC AAD56197;
 XX
 DT 07-AUG-2003 (first entry)
 XX
 DE Human LRRCAPS DNA #11.
 XX
 KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 XX LRRCAPS; cancer; gene therapy; ds.
 OS Homo sapiens.
 XX

PN WO2003035831-A2.
 PD 01-MAY-2003.
 XX
 PF 21-OCT-2002; 2002WO-US033540.
 XX
 PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX
 XX WPI; 2003-421410/39.
 DR
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 PS Example 5; Page 73-74; 99pp; English.
 XX
 CC The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified Leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 167 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-6 (1-9) x AAD56197 (1-2053)
 Qy 1 AlaLeuileGlyAlailePheLeuLeu 9
 Db 1174 GCCCTGATAGGCGCTATTTCCTCTCG 1200
 RESULT 22
 AAD56200
 ID AAD56200 standard; DNA; 2053 BP.
 XX
 AC AAD56200;
 XX
 DT 07-AUG-2003 (first entry)
 XX
 DE Human LRRCAPS DNA #12.
 XX
 KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003035831-A2.
 PD 01-MAY-2003.
 XX
 PF 21-OCT-2002; 2002WO-US033540.
 XX
 PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 XX

```
PR 01-MAR-2002; 2002US-0361196P.
XX (EXEL-) EXELIXIS INC.
PA
PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PT Francis-Lang H, Friedman L;
XX WPI; 2003-421410/39.
DR
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
PS Disclosure; Page 76-77; 99pp; English.
XX
CC The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS DNA
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 167 Length: 2053
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x AAD56200 (1-2053)
Qy 1 AlaLeuIlleGlyAlaIlePheLeuLeu 9
Db 1174 GCCCTGATAGGCGCTATTTCCTCTG 1200

RESULT 23
ADN38721
ID ADN38721 standard; cDNA; 2053 BP.
XX
AC ADN38721;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397759P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevesi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
XX P-PSDB; ADN38722.
XX
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX
XX Claim 9; SEQ ID NO 39; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
```

KW cancer; cytostatic; gene; ss.
 XX Homo sapiens.
 OS
 XX WO2004016225-A2.
 PN
 XX 26-FEB-2004.
 PD
 XX 19-AUG-2003; 2003WO-US025892.
 PF
 XX 19-AUG-2002; 2002US-0404809P.
 PR 21-AUG-2002; 2002US-0405645P.
 PR 23-SEP-2002; 2002US-0413192P.
 PR 15-OCT-2002; 2002US-0419008P.
 PR 15-NOV-2002; 2002US-0426847P.
 PR 02-JUL-2003; 2003US-0484959P.
 XX (GETH) GENENTECH INC.
 PA
 XX Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
 PI Spencer SD, Wu TD, Zhang Z;
 PI
 XX WPI; 2004-257144/24.
 DR P-PSDB; ADL06552.
 DR
 XX New antibody that binds to a tumor-associated antigenic target (TAT)
 PT polypeptide, useful for preparing a composition for diagnosing or
 PT treating cancer.
 PT
 XX Claim 1; SEQ ID NO 53; 319pp; English.
 PS
 XX The present invention relates to the isolation of human tumour-associated
 CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also
 CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is
 CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a
 CC humanised antibody. It is conjugated to a growth inhibitory agent. It is
 CC produced in bacteria or in CHO cells and induces death of a cell to which
 CC it binds. The antibody is useful for preparing a composition for
 CC diagnosing or treating tumours and cancer. The present sequence
 CC represents a human TAT cDNA sequence of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 167 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-6 (1-9) x ADL06473 (1-2053)
 QY 1 AlaLeuileGlyAlaIlePheLeuLeu 9
 DB 1174 GCCCTGATAGGCGCTATTTCTCTCTG 1200

RESULT 25
 ADN03961
 ID ADN03961 standard; cDNA; 2053 BP.
 XX
 AC ADN03961;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Antipsoriatic cDNA sequence #180.
 XX
 KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
 XX Homo sapiens.
 OS
 XX WO2004028479-A2.
 PN

PD 08-APR-2004.
 XX
 PF 25-SEP-2003; 2003WO-US030907.
 XX
 PR 25-SEP-2002; 2002US-0414006P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 PI Wu TD;
 PI
 XX WPI; 2004-305105/28.
 DR P-PSDB; ADN03962.
 DR
 XX New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 PT
 XX Claim 1; SEQ ID NO 355; 3069pp; English.
 PS
 XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polynucleotides of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 167 Length: 2053
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-6 (1-9) x ADN03961 (1-2053)
 QY 1 AlaLeuileGlyAlaIlePheLeuLeu 9
 DB 1174 GCCCTGATAGGCGCTATTTCTCTCTG 1200

RESULT 26
 ADN25444
 ID ADN25444 standard; DNA; 2053 BP.
 XX
 AC ADN25444;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Breast cancer prognosis marker #1305.
 XX
 KW ds; breast cancer; prognosis; gene expression; diagnosis.
 XX Homo sapiens.
 OS
 XX WO2004065545-A2.
 PN
 XX
 PD 05-AUG-2004.
 XX
 PF 15-JAN-2004; 2004WO-US001100.
 XX
 PR 15-JAN-2003; 2003US-00342887.
 XX
 PA (ROSE-) ROSETTA INPHARMATICS LLC.
 PA (NECA-) NETHERLANDS CANCER INST.
 XX
 XX Van't Veer LJ, He Y;
 PI
 XX WPI; 2004-593473/57.
 DR
 XX
 PT Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of

PT five genes in a cell sample taken from patient, to control levels.
XX Disclosure; SEQ ID NO 1305; 226pp; English.
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 167 Length: 2053
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-6 (1-9) x ADR25444 (1-2053)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9
Db 1174 GCCCTGATAGGCGCTATTTCCTCTG 1200

RESULT 27
ACN38510
ID ACN38510 standard; cDNA; 2053 BP.
XX
AC ACN38510;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA103471, SEQ ID NO:2070.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
WO2004030615-A2.
XX
PN
XX
PD 15-APR-2004.
XX
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
XX (GETH) GENENTECH INC.
XX
XX Wu TD, Zhang Z, Zhou Y;
PI
XX
XX WPI; 2004-347921/32.
XX
XX P-PSDB; ABM80804.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
XX useful in preparing a medicament for treating or detecting a
XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX prostate cancer or tumor.
XX
XX Claim 1; SEQ ID NO 2070; 7273pp; English.
XX

CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 167 Length: 2053
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-6 (1-9) x ACN38510 (1-2053)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9
Db 1174 GCCCTGATAGGCGCTATTTCCTCTG 1200

RESULT 28
ADV35098
ID ADV35098 standard; cDNA; 2053 BP.
XX
AC ADV35098;
XX
DT 10-FEB-2005 (first entry)
XX
DE Human cDNA of an exemplary efficacy gene for BAD SeqID174.
XX
KW human; ss; multi-parameter high throughput screening; MPHTS;
KW disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
KW bipolar affective disorder; BAD; autism; Parkinson's;
KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.
XX
OS Homo sapiens.
XX
XX
XX US2003096264-A1.
XX
XX 22-MAY-2003.
XX
XX 18-JUN-2002; 2002US-00175523.
XX
XX 18-JUN-2001; 2001US-0299151P.
XX 07-SEP-2001; 2001US-0317828P.
XX 25-SEP-2001; 2001US-0325150P.
XX 14-NOV-2001; 2001US-0333047P.
XX 18-JAN-2002; 2002US-0349936P.
XX 04-MAR-2002; 2002US-0361834P.
XX
XX (PSYC-) PSYCHIATRIC GENOMICS INC.
XX
XX Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
XX Palfreyman M, Rajan P;
XX WPI; 2004-118903/12.
XX

XX Identifying a compound that can treat disease or disorders, such as, a
PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
PT determining the expression of one or more efficacy genes in a cell
PT contacted with the test compound.
XX
PS Example 6; SEQ ID NO 174; 39pp; English.
XX
CC This invention relates to a novel screening method identified as a multi-
CC parameter high throughput screening (MPHTS) assay. Specifically, it
CC refers to an assay that utilizes the disease signature of a plurality of
CC specific genes associated with a particular disease, and identifies
CC differential expression between those cells taken from individuals
CC affected by that disease and those that are not affected. The present
CC invention then describes the screening of candidate pharmaceutical
CC compounds to identify those that have a potential therapeutic benefit for
CC the treatment of neuropsychiatric and neurodegenerative disorders
CC including schizophrenia, bipolar affective disorder (BAD) and autism, as
CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
CC of this invention exhibit various activities including neuroleptic,
CC nootropic, antimanic and antidepressant. Furthermore, the screening
CC method used in MPHTS will be automated, such that a large number of test
CC compounds may be rapidly screened with a minimal amount of labour and
CC effort. This polynucleotide is a human cDNA sequence of a gene that is
CC differentially expressed in the presence of a therapeutic compound and
CC represents an exemplary efficacy gene for bipolar affective disorder,
CC given in an exemplification of the invention.
XX
SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 167 Length: 2053
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-6 (1-9) x ADV35098 (1-2053)

QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
DB 1174 GCCCTGATAGGCGCTATTTCCTCTG 1200

RESULT 29
AAS87175
ID AAS87175 standard; cDNA; 2338 BP.

XX AAS87175;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #22979.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSEB-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI, 2001-639362/73.

DR P-PSDB; ABG22988.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX

PS Claim 1; SEQ ID NO 22979; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 193 Length: 2338
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-6 (1-9) x AAS87175 (1-2338)

QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
DB 1431 GCCCTGATAGGCGCTATTTCCTCTG 1457

RESULT 30
AAK94253
ID AAK94253 standard; cDNA; 2359 BP.

XX AAK94253;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human full-length cDNA, SEQ ID NO: 2864.
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX
XX 11-JAN-2000; 2000JP-00118774.
XX
XX 02-MAY-2000; 2000JP-00183765.
XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Iehii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR P-PSDB; AAM93333.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX
 SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 195 Length: 2359
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-6 (1-9) x AAK94253 (1-2359)
 QY 1 AlaLeuileGlyAlaIlePheLeuLeu 9
 DB 1513 GCCCTGATAGGCGCTATTTCTCTCTG 1539
 RESULT 31
 ADL30831
 ID ADL30831 standard; cDNA; 2359 BP.
 XX
 AC ADL30831;
 DT 20-MAY-2004 (first entry)
 XX
 DE Full length human cDNA clone SeqID 2864.
 XX human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gens.
 XX Homo sapiens.
 XX EP1396543-A2.
 XX 10-MAR-2004.
 XX
 XX 07-JUL-2000; 2003EP-00025638.
 XX
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Iehii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2004-204755/20.
 DR P-PSDB; ADL30832.
 DR

XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX
 XX Example 1; SEQ ID NO 2864; 1340pp; English.
 PS
 CC This invention relates to a novel primers useful for synthesising full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.
 XX
 SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 195 Length: 2359
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-6 (1-9) x ADL30831 (1-2359)
 QY 1 AlaLeuileGlyAlaIlePheLeuLeu 9
 DB 1513 GCCCTGATAGGCGCTATTTCTCTCTG 1539
 RESULT 32
 AAK94254
 ID AAK94254 standard; cDNA; 2361 BP.
 XX
 AC AAK94254;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human full-length cDNA, SEQ ID NO: 2866.
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX Homo sapiens.
 XX EP1130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-00114089.
 XX
 PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 XX Ota T, Nishikawa T, Isogai T, Hayashi K, Iehii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR P-PSDB; AAM93334.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.
 CC The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX

CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO

SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	196	Length:	2361
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	4	Gaps:	0

US-10-774-176-6 (1-9) x AAK94254 (1-2361)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9

DB 1515 GCCCTGATAGGCGCTATTTCCTCTG 1541

RESULT 33

AD126162
 ID AD126162 standard; cDNA; 2361 BP.

XX AC AD126162;

XX DT 22-APR-2004 (first entry)

XX DE Human cDNA encoding protein that promotes STAT6 activation #64.

XX KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

XX OS Homo sapiens.

XX PN WO2003104277-A2.

XX PD 18-DEC-2003.

XX PF 05-JUN-2003; 2003WO-JP007123.

XX PR 05-JUN-2002; 2002JP-00164257.

XX PR 06-JUN-2002; 2002US-0385912P.

XX PR 26-DEC-2002; 2002JP-00377326.

XX PR 27-DEC-2002; 2002US-0436467P.

XX PR 15-MAY-2003; 2003JP-00137505.

XX PR 16-MAY-2003; 2003US-0470836P.

XX PA (ASAH) ASAH KASEI KK.

XX PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX WPI; 2004-122214/12.

XX P-PSDB; ADI26163.

XX PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.

XX PS Claim 4; SEQ ID NO 127; 1368pp; English.

XX CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is

CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.

SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	196	Length:	2361
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-6 (1-9) x ADI26162 (1-2361)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9

DB 1515 GCCCTGATAGGCGCTATTTCCTCTG 1541

RESULT 34

AD130833

ID AD130833 standard; cDNA; 2361 BP.

XX AC AD130833;

XX DT 20-MAY-2004 (first entry)

XX DE Full length human cDNA clone SeqID 2866.

XX KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.

XX OS Homo sapiens.

XX PN EP1396543-A2.

XX PD 10-MAR-2004.

XX PF 07-JUL-2000; 2003EP-00025638.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JAN-2000; 2000JP-00118774.

XX PR 02-MAY-2000; 2000JP-00183865.

XX PR 07-JUL-2000; 2000EP-00114089.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2004-204755/20.

XX P-PSDB; ADL30834.

XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.

XX Example 1, SEQ ID NO 2866; 1340pp; English.

XX This invention relates to a novel primers useful for synthesizing full

CC length cDNA molecules that encode human proteins. Specifically, it refers

CC to secretory or membrane proteins that are potential therapeutic agents/

CC target molecules in the field of medicine, and in particular genes

CC encoding proteins that are associated with signal transduction,

CC glycoproteins and transcription. The present invention describes a method

CC for efficiently cloning a full length human cDNA from both the 5' and 3'

CC ends using the oligo-capping method. This polynucleotide sequence is a

CC full length human cDNA clone of the invention.

XX

SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	196	Length:	2361
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-6 (1-9) x ADL30833 (1-2361)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9

Db 1515 GCCCTGATAGCGCTATTTTCTCTG 1541

RESULT 35

ADI26160

ID ADI26160 standard; cDNA; 2557 BP.

XX AC ADI26160;

XX

DT 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #63.

XX

ss; gene; human; signal transducer and activator of transcription 6;

KW STAT6; immunogen; STAT6 activation; allergy; inflammation;

KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;

KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;

KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

XX

OS Homo sapiens.

XX

PN WO2003104277-A2.

XX

PD 18-DEC-2003.

XX

05-JUN-2003; 2003WO-JP007123.

XX

05-JUN-2002; 2002JP-00164257.

XX

06-JUN-2002; 2002US-0385912P.

XX

26-DEC-2002; 2002JP-00377326.

XX

27-DEC-2002; 2002US-0436467P.

XX

15-MAY-2003; 2003JP-00137505.

XX

16-MAY-2003; 2003US-0470836P.

XX

(ASAH) ASAHI KASEI KK.

XX

PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX

WPI; 2004-122214/12.

XX

P-P8DB; ADI26161.

XX

New signal transducer and activator of transcription 6 activation

PT promoting purified protein, for diagnosing and treating disease

PT associated with activation/inhibition of transcription factor e.g.

PT diabetes and cancer.

PT

PS Claim 4; SEQ ID NO 125; 1368pp; English.

XX

The invention relates to a purified protein promoting signal transducer

CC and activator of transcription 6 activation (STAT6). The protein is

CC useful for the producing an antibody, which involves administering the

CC protein or its epitope-bearing fragments to a non-human animal as an

CC antigen. The nucleic acid is useful for diagnosing a disease or

CC susceptibility to a disease related to expression or activity of the

CC protein. A transformant expressing the protein is useful for screening

CC compounds which inhibit or promote STAT6 activation. A transformant

CC expressing the protein is useful for producing a pharmaceutical

CC composition. Compositions, antibodies and antisense molecules are useful

CC for the treating a disease associated with STAT6 activation such as

CC allergic diseases, inflammation, autoimmune diseases, diabetes,

CC hyperlipidaemia, infections disease and cancers. Compositions are useful

CC for treating disease associated with STAT6 activation and/or prevention

CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid

CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,

CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,

CC viral hepatitis and AIDS. The protein has efficient promoting STAT6

CC activity. The protein or nucleic acid is effectively useful for screening

CC compounds for treating and preventing disease associated with excessive

CC activation or inhibition of STAT6. The present sequence represents a

CC human cDNA encoding a protein which promotes STAT6 activation.

XX

SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	214	Length:	2557
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-6 (1-9) x ADI26160 (1-2557)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9

Db 1663 GCTCTGATAGCGCTATTTTCTCTC 1689

RESULT 36

ADI26158

ID ADI26158 standard; cDNA; 2557 BP.

XX AC ADI26158;

XX

DT 22-APR-2004 (first entry)

XX Human cDNA encoding protein that promotes STAT6 activation #62.

XX

ss; gene; human; signal transducer and activator of transcription 6;

KW STAT6; immunogen; STAT6 activation; allergy; inflammation;

KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;

KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;

KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;

KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

XX

OS Homo sapiens.

XX

PN WO2003104277-A2.

XX

PD 18-DEC-2003.

XX

05-JUN-2003; 2003WO-JP007123.

XX

05-JUN-2002; 2002JP-00164257.

XX

06-JUN-2002; 2002US-0385912P.

XX

26-DEC-2002; 2002JP-00377326.

XX

27-DEC-2002; 2002US-0436467P.

XX

15-MAY-2003; 2003JP-00137505.

XX

16-MAY-2003; 2003US-0470836P.

XX

(ASAH) ASAHI KASEI KK.

XX

PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;

XX

WPI; 2004-122214/12.

XX

P-P8DB; ADI26161.

XX

New signal transducer and activator of transcription 6 activation

PT promoting purified protein, for diagnosing and treating disease

PT associated with activation/inhibition of transcription factor e.g.

PT diabetes and cancer.

PT

PA (ASAH) ASahi KASEI KK.
XX Sugahara T, Mateuda A, Honda G, Muramatsu S, Ishizawa K;
PI WPI; 2004-122214/12.
XX P-PSDB; ADI26159.
DR New signal transducer and activator of transcription 6 activation
DR promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX Claim 4; SEQ ID NO 123; 1368pp; English.
XX The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.
XX XX
XX Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 214 Length: 2557
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0
US-10-774-176-6 (1-9) x ADI26158 (1-2557)
Qy 1 AlaLeulleGlyAlaIlePheLeuleu 9
Db 1663 GCTCTGATAGGCGCTATTTCCTCCTC 1689
RESULT 37
ADF02237
ID ADF02237 standard; DNA; 1494 BP.
XX AC ADF02237;
XX DT 12-FEB-2004 (first entry)
XX XX
XX Bacterial polynucleotide #2522.
XX Proteus mirabilis infection; bacterial infection; antibacterial;
XX immunostimulant; gene; ds.
XX Proteus mirabilis.
XX OS
XX US6605709-B1.
XX PN
XX 12-AUG-2003.
XX PD
XX 05-APR-2000; 2000US-00543681.
XX PF

XX 09-APR-1999; 99US-0128706P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL;
XX WPI; 2003-995291/82.
DR P-PSDB; ADF06409.
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX Disclosure; SEQ ID NO 2522; 870pp; English.
XX The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polynucleotide of the invention.
XX XX
XX Sequence 1494 BP; 356 A; 259 C; 342 G; 537 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 190 Length: 1494
Score: 39.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 97.5% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-6 (1-9) x ADF02237 (1-1494)
Qy 1 AlaLeulleGlyAlaIlePheLeuleu 9
Db 535 GCCCTGATTGGTGGCGTATTCTTATTA 561
RESULT 38
ABN60581/c
ID ABN60581 standard; cDNA; 433 BP.
XX AC ABN60581;
XX DT 28-JUN-2002 (first entry)
XX XX
XX Human cancer related polynucleotide SEQ ID NO 548.
XX Human; cytostatic; gene expression; gene mapping; tissue profiling;
XX gene therapy; cancer; tumour; gene; ss.
XX Homo sapiens.
XX OS
XX WO200214500-A2.
XX PN
XX 21-FEB-2002.
XX PD
XX 16-AUG-2001; 2001WO-US025840.
XX PF
XX 16-AUG-2000; 2000US-0226326P.
XX PR
XX (CHIR) CHIRON CORP.
XX PA (HYSE-) HYSEQ INC.
XX XX
XX Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;

PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
 XX WPI; 2002-241905/29.
 XX New nucleic acid for producing a polypeptide, detecting differentially
 PT expressed genes correlated with a cancerous state of a mammalian cell,
 PT and inhibiting tumor growth.
 XX
 XX Claim 1; SEQ ID NO 548; 883pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
 CC with cytostatic activity. The polynucleotide is used to produce a
 CC polypeptide, to detect differentially expressed genes correlated with a
 CC cancerous state of a mammalian cell and to inhibit tumour growth. The
 CC polynucleotide is used as a probe in mapping and tissue profiling. The
 CC encoded polypeptide and antibodies to the polypeptide can also be used
 CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
 CC gene therapy. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 433 BP; 134 A; 93 C; 155 G; 51 T; 0 U; 0 Other;
 Alignment Scores: 76.1 Length: 433
 Pred. No.: 38.00 Matches: 8
 Score: 100.0% Conservative: 1
 Percent Similarity: 88.9% Mismatches: 0
 Best Local Similarity: 88.9% Indels: 0
 Query Match: 95.0% Gaps: 0
 DB: 6
 US-10-774-176-6 (1-9) x ABN60581 (1-433)
 QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
 DB 200 GCTCTTCTAGGCGCCATCTTCTCTC 174
 RESULT 39
 ABN60925/c
 ID ABN60925 standard; cDNA; 567 BP.
 AC ABN60925;
 XX
 XX 28-JUN-2002 (first entry)
 DT
 XX Human cancer related polynucleotide SEQ ID NO 892.
 DE
 XX Human; cytostatic; gene expression; gene mapping; tissue profiling;
 KW gene therapy; cancer; tumour; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200214500-A2.
 PN
 XX 21-FEB-2002.
 PD
 XX 16-AUG-2001; 2001WO-US025840.
 XX
 XX 16-AUG-2000; 2000US-0236326P.
 XX
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 XX Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
 PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
 PI WPI; 2002-241905/29.
 XX
 XX New nucleic acid for producing a polypeptide, detecting differentially
 PT expressed genes correlated with a cancerous state of a mammalian cell,
 PT and inhibiting tumor growth.
 XX
 XX Claim 1; SEQ ID NO 892; 883pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
 CC with cytostatic activity. The polynucleotide is used to produce a
 CC polypeptide, to detect differentially expressed genes correlated with a
 CC cancerous state of a mammalian cell and to inhibit tumour growth. The
 CC polynucleotide is used as a probe in mapping and tissue profiling. The
 CC encoded polypeptide and antibodies to the polypeptide can also be used
 CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
 CC gene therapy. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 567 BP; 171 A; 135 C; 195 G; 66 T; 0 U; 0 Other;
 Alignment Scores: 103 Length: 567
 Pred. No.: 38.00 Matches: 8
 Score: 100.0% Conservative: 1
 Percent Similarity: 88.9% Mismatches: 0
 Best Local Similarity: 88.9% Indels: 0
 Query Match: 95.0% Gaps: 0
 DB: 6
 US-10-774-176-6 (1-9) x ABN60925 (1-567)
 QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
 DB 188 GCTCTTCTAGGCGCCATCTTCTCTC 162
 RESULT 40
 ADO35702
 ID ADO35702 standard; DNA; 1101 BP.
 XX
 AC ADO35702;
 XX
 XX 26-AUG-2004 (first entry)
 DT
 XX Novel mouse gene sequence #375.
 DE
 XX mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
 KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
 KW viral disorder; ds; gene.
 XX
 XX Mus sp.
 OS
 XX WO2004046310-A2.
 PN
 XX 03-JUN-2004.
 PD
 XX 24-OCT-2003; 2003WO-US033948.
 XX
 XX 15-NOV-2002; 2002US-0426916P.
 PR
 XX 04-DEC-2002; 2002US-0431158P.
 PR
 XX 05-DEC-2002; 2002US-0431445P.
 PR
 XX 05-DEC-2002; 2002US-0431606P.
 PR
 XX 09-JUN-2003; 2003US-0476621P.
 PR
 XX 09-JUN-2003; 2003US-0476632P.
 PR
 XX 08-JUL-2003; 2003US-0485217P.
 PR
 XX 08-JUL-2003; 2003US-0485359P.
 PR
 XX 08-AUG-2003; 2003US-0493332P.
 PR
 XX 08-AUG-2003; 2003US-0493356P.
 XX
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
 PA
 XX Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
 PI WPI; 2004-431966/40.
 XX
 XX New mouse nucleic acid molecules and polypeptides, useful for treating
 PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
 PT disease or thrombosis.
 XX
 XX Claim 1; SEQ ID NO 375; 263pp; English.

CC The invention comprises 744 novel mouse DNA sequences (genes). The DNA
 CC sequences of the invention are useful for treating cancer, psoriasis,
 CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
 CC immune disorders, bacterial disorders and viral disorders. The present
 CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
 CC present DNA sequence is not shown in the specification, but has been
 CC retrieved from the WIPO website.

SQ Sequence 1101 BP; 216 A; 332 C; 355 G; 198 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 220 Length: 1101
 Score: 38.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-6 (1-9) x ADO35702 (1-1101)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
 Db 394 GCCCTGCTCGAGCCATTTCCTC 420

RESULT 41

ACA28481
 ID ACA28481 standard; DNA; 951 BP.

AC ACA28481;

XX 19-JUN-2003 (first entry)

DT Prokaryotic essential gene #10138.

XX

DE Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.

XX

OS Clostridium botulinum.

XX

FN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

XX

PR 06-SEP-2001; 2001US-00948993.

XX

PR 25-OCT-2001; 2001US-0342923P.

XX

PR 08-FEB-2002; 2002US-00072851.

XX

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;

XX

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR P-PSDB; ABU24611.

XX

DR WPI; 2003-029926/02.

XX

DR P-PSDB; ABU24611.

XX

DR New antisense nucleic acids, useful for identifying proteins or screening

XX

PT for homologous nucleic acids required for cellular proliferation to

XX

PT isolate candidate molecules for rational drug discovery programs.

XX

XX Claim 14; SEQ ID NO 16351; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of

XX

CC the 6213 antisense sequences given in the specification where expression

XX

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX

CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX

CC encoding a polypeptide whose expression is inhibited by the antisense

XX

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX

CC polypeptide or its fragment whose expression is inhibited by the

XX

CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 951 BP; 319 A; 107 C; 160 G; 365 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 305 Length: 951
 Score: 37.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 92.5% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x ACA28481 (1-951)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9

Db 802 GCATTGCTTGAGCAGTATTCTCTTA 828

RESULT 42

ABN70062

ID ABN70062 standard; DNA; 225 BP.

XX

AC ABN70062;

XX

DT 01-JUL-2002 (first entry)

XX

DE Streptococcus polynucleotide SEQ ID NO 8037.

XX

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX

KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;

XX

OS Streptococcus pyogenes.

XX

PN WO200234771-A2.

XX

PD 02-MAY-2002.

XX

PF 29-OCT-2001; 2001WO-GB004789.

XX

PR 27-OCT-2000; 2000GB-00026333.

XX

PR 24-NOV-2000; 2000GB-00028727.

XX

PR 07-MAR-2001; 2001GB-00005640.

XX

(CHIR-) CHIRON SPA.

XX

PA (GENO-) INST GENOMIC RES.

XX

PI Telford J, Maignani V, Margarit Y Rosl, Grandi G, Fraser C;

XX

PI Tettelin H;

XX

DR WPI; 2002-352536/38.

DR P-PSDB; ABG22987.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 22978; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have application in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 453 BP; 108 A; 111 C; 113 G; 121 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.: 215 Length: 453
Score: 36.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-6 (1-9) x AAS87174 (1-453)
XX

Qy 1 AlaLeuileGlyAlailePheLeu 8
Db 277 GCCCTGATAGCGCTATTTCCTC 300

RESULT 45
ACF71997
ID ACF71997 standard; DNA; 681 BP.
XX
XX ACF71997;
AC
XX
DT 20-NOV-2003 (first entry)
XX
XX
DE Photorhabdus luminescens nucleotide sequence #10464.
XX
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough; gene; ds.
XX
XX Photorhabdus luminescens.
OS
XX
XX WO200294867-A2.
PN
XX
XX 28-NOV-2002.
PD
XX
XX 07-FEB-2002; 2002WO-IB003040.
PF
XX
XX 07-FEB-2001; 2001FR-00001659.
PR
XX
XX (INSP) INST PASTEUR.
PA

PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX
XX WPI; 2003-148459/14.
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
PT
XX Claim 2; SEQ ID NO 10464; 1205pp; French.
XX
CC The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
SQ Sequence 681 BP; 183 A; 135 C; 180 G; 183 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.: 341 Length: 681
Score: 36.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-6 (1-9) x ACF71997 (1-681)
XX

Qy 1 AlaLeuileGlyAlailePheLeu 8
Db 352 GCGCTGATGGTCTATTCCTT 375

RESULT 46
ADF03189
ID ADF03189 standard; DNA; 687 BP.
XX
XX ADF03189;
AC
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX Bacterial polynucleotide #3474.
DE
XX
XX Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant; gene; ds.
XX
XX Proteus mirabilis.
OS
XX
XX US6605709-B1.
PN
XX
XX 12-AUG-2003.
PD
XX
XX 05-APR-2000; 2000US-00543681.
PF
XX
XX 09-APR-1999; 99US-0128706P.
PR
XX
XX

(GENO-) GENOME THERAPEUTICS CORP.
 PA Breton GL;
 PI WPI; 2003-895291/82.
 XX P-PSDB; ADF07361.
 DR New Proteus mirabilis polypeptides and polynucleotides, useful as
 XX reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX
 XX Disclosure; SEQ ID NO 3474; 870pp; English.
 XX
 CC The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunising an individual against P. mirabilis, a
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polynucleotide of the invention.
 XX
 SQ Sequence 687 BP; 211 A; 127 C; 152 G; 137 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 345 Length: 687
 Score: 36.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 90.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-6 (1-9) x ADF03189 (1-687)
 QY 1 AlaLeuileGlyAlaIlePheLeu 8
 DB 358 GCGTAAATGGTGTATTTCTT 381
 RESULT 47
 AAC92268
 ID AAC92268 standard; DNA; 922 BP.
 XX
 AC AAC92268;
 DT 22-MAR-2001 (first entry)
 XX
 XX Green fluorescent protein (GFP) encoding nucleotide sequence.
 XX
 XX Caspase; caspase cleavage site; green fluorescent protein; GFP; FRFT;
 KW fluorescent resonance energy transfer; fusion protein; identification;
 KW caspase inhibitor; apoptosis; neurodegenerative disease; stroke;
 KW Alzheimer's disease; Parkinson's disease; traumatic brain injury;
 KW ischaemia-reperfusion injury; graft-versus-host disease;
 KW autoimmune disorder; inflammatory disease; neuronal damage reduction;
 KW bacterial meningitis; ds.
 XX
 OS Unidentified.
 XX
 XX WO200073437-A1.
 PN 07-DEC-2000.
 PD
 XX 25-MAY-2000; 2000WO-CA000620.
 XX
 XX 27-MAY-1999; 99US-0136286P.
 XX
 XX (MERI) MERCK FROSST CANADA & CO.
 PA

PI Xanthoudakis S, Tawa P, Caseady R, Nicholson D;
 XX WPI; 2001-070966/08.
 DR P-PSDB; AAB51230.
 XX
 PT New fusion protein comprising an ultra-bright donor and an acceptor green
 PT fluorescent protein linked by a peptide having at least one caspase
 PT cleavage site, for identifying activators or inhibitors of caspase in
 PT living cells or in vitro.
 XX
 PS Disclosure; Fig 7A; 50pp; English.
 XX
 CC The present invention describes a fusion protein (I) comprising of a
 CC donor green fluorescent protein (GFP) that is an ultra-bright GFP and an
 CC acceptor GFP linked by a peptide comprising at least one caspase cleavage
 CC site. AAB51201 to AAB51225 represent specifically claimed caspase
 CC cleavage site peptide sequences from the present invention. The fusion
 CC protein is useful in fluorescent resonance energy transfer (FRFT) and in
 CC the identification of activators or inhibitors of caspase activity in
 CC living cells or in vitro. Methods from the present invention can be used
 CC in the identification of caspase inhibitors which can be used in treating
 CC apoptosis, neurodegenerative diseases such as Alzheimer's disease and
 CC Parkinson's disease, traumatic brain injury, stroke, ischaemia-
 CC reperfusion injury, graft-versus-host disease and autoimmune disorders.
 CC These caspase inhibitors may also be used in clinical trial for treatment
 CC of inflammatory diseases or reduction of neuronal damage in a rabbit
 CC model of bacterial meningitis. The present sequence encodes a GFP which
 CC is used in the exemplification of the present invention
 XX
 SQ Sequence 922 BP; 322 A; 164 C; 170 G; 266 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 482 Length: 922
 Score: 36.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 90.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-6 (1-9) x AAC92268 (1-922)
 QY 2 LeuileGlyAlaIlePheLeu 9
 DB 879 TTATAGGGCTATTTCTTATTA 902
 RESULT 48
 AAL57769
 ID AAL57769 standard; DNA; 922 BP.
 XX
 AC AAL57769;
 DT 06-NOV-2003 (first entry)
 XX
 XX Aequorea victoria green fluorescent protein (GFP) DNA sequence.
 XX
 XX Locus specific targeting fragment; eukaryotic cell;
 KW homologous recombination; endogenous mismatch repair; gene function;
 KW gene expression; over-producer clone; high protein production;
 KW therapeutic target discovery; pharmacological compound screening;
 KW protein manufacturing; mismatch repair protein;
 KW green fluorescent protein; GFP; gene; ds.
 XX
 OS Aequorea victoria.
 XX
 XX Key Location/Qualifiers
 FT CDS 26..742
 FT /*tag= a
 FT /product= "Aequorea victoria GFP"
 XX
 XX WO2003062435-A1.
 XX
 XX 31-JUL-2003.
 PD

PF 17-JAN-2003; 2003WO-US001361.
XX
PR 18-JAN-2002; 2002US-034956SP.
XX
PA (MORP-) MORPHOTEK INC.
XX
PI Grasso L, Kline JB, Nicolaides NC, Sass PW;
XX
PI WPI; 2003-646070/61.
XX
DR P-PSDB; AAO27520.
XX
PT Introducing a locus specific targeting fragment into the genome of a cell
PT through homologous recombination by inhibiting endogenous mismatch repair
PT of the cell and introducing a locus specific targeting fragment into the
PT cell.
XX
PS Disclosure; Page 90-91; 110pp; English.
XX
CC This invention relates to a novel method for the introduction of a locus
CC specific targeting fragment into the genome of a eukaryotic cell through
CC homologous recombination and comprises the inhibition of endogenous
CC mismatch repair of the cell. The locus specific targeting fragment is a
CC polynucleotide comprising at least one promoter, a selectable marker and
CC 5' and 3' flanking regions of 20-120 nucleotides. The 5' and 3' flanking
CC regions are homologous to a selected portion of the genome of the cell.
CC The method of the invention may be useful for introducing a locus
CC specific targeting fragment into the genome of a cell through homologous
CC recombination. This may be useful for studying gene function, gene
CC expression and generating over-producer clones for high protein
CC production. In addition, the invention may be useful for therapeutic
CC target discovery, pharmacological compound screening and protein
CC manufacturing. The present sequence is the DNA sequence of Aequorea
CC victoria green fluorescent protein (GFP) which was used as a reporter
CC element during the exemplification of the method of the invention
XX
SQ Sequence 922 BP; 322 A; 164 C; 170 G; 266 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 482 Length: 922
Score: 36.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-6 (1-9) x AAL57769 (1-922)

QY 2 LeuileGlyAlaIlePheLeuLeu 9
DB 879 TTAATAGGGGCTATTTCTTATTA 902

RESULT 49
ABL05253
ID ABL05253 standard; cDNA; 3222 BP.
XX
AC ABL05253;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10241.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX

PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
PI WPI; 2001-656860/75.
XX
DR P-PSDB; ABB61150.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 10241; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3222 BP; 800 A; 823 C; 837 G; 762 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 26+03 Length: 3222
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-6 (1-9) x ABL05253 (1-3222)

QY 1 AlaLeuileGlyAlaIlePheLeuLeu 9
DB 2240 GCTCTGCTGGTGCCATATTCCTCATC 2266

RESULT 50
ABL05252
ID ABL05252 standard; cDNA; 8737 BP.
XX
AC ABL05252;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10238.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
PI WPI; 2001-656860/75.
XX
DR P-PSDB; ABB61149.
XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT Genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 10238; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 8737 BP; 2394 A; 1869 C; 1894 G; 2580 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.22e+03 Length: 8737
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-6 (1-9) x ABL05252 (1-8737)

Qy 1 AlaLeuileGlyAlaIlePheLeu 9
|||||:|||||:|||||:|||||:|
Db 6574 GCTCTGCTGGTGCCATATCTTCATC 6600

Search completed: April 25, 2006, 12:37:14
Job time : 323.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignments)
171.290 Million cell updates/sec

Title: US-10-774-176-6

Perfect score: 40

Sequence: 1 ALIGAIFLL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5893141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSWEB/spool/US10774176/runat_24042006.165114.19197/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUPERIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-OUTFMT=ptc -NORM=score -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-USER=US10774176 @CCN 1 1 6765 @runat_24042006.165114.19197 -NCPU=6 -ICPU=3
-NO_MMAP -NRG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -PSLEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sv.*
12: gb.un.*
13: gb.vi.*
14: gb.htg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	290	6	CQ687716 Sequence
2	40	100.0	475	6	CQ920916 Sequence
3	40	100.0	901	6	BD249733 Polypepti

4	40	100.0	901	6	AX025013	AX025013 Sequence
5	40	100.0	901	6	AX316088	AX316088 Sequence
6	40	100.0	927	6	AX829164	AX829164 Sequence
7	40	100.0	1260	6	AX467373	AX467373 Sequence
8	40	100.0	1260	6	AX821533	AX821533 Sequence
9	40	100.0	1260	6	AX821548	AX821548 Sequence
10	40	100.0	1263	6	BD249731	BD249731 Polypepti
11	40	100.0	1263	6	AX025011	AX025011 Sequence
12	40	100.0	1263	6	AX149553	AX149553 Sequence
13	40	100.0	1263	6	AX316086	AX316086 Sequence
14	40	100.0	1263	6	AX467371	AX467371 Sequence
15	40	100.0	1281	6	BD249732	BD249732 Polypepti
16	40	100.0	1281	6	AX025012	AX025012 Sequence
17	40	100.0	1281	6	AX316087	AX316087 Sequence
18	40	100.0	2053	6	CQ731678	CQ731678 Sequence
19	40	100.0	2053	8	HS5740A	Z29083 Homo sapien
20	40	100.0	2333	9	AF063939	AF063939 Rattus no
21	40	100.0	2359	6	BD127282	BD127282 Primer fo
22	40	100.0	2359	6	CQ782724	CQ782724 Sequence
23	40	100.0	2359	8	AK074786	AK074786 Homo sapi
24	40	100.0	2361	6	BD127283	BD127283 Primer fo
25	40	100.0	2361	6	CQ782726	CQ782726 Sequence
26	40	100.0	2361	8	AX961916	AX961916 Sequence
27	40	100.0	2361	8	AK074790	AK074790 Homo sapi
28	40	100.0	2361	9	BC087011	BC087011 Rattus no
29	40	100.0	2379	8	BC037161	BC037161 Homo sapi
30	40	100.0	2423	9	BC058198	BC058198 Mus muscu
31	40	100.0	2557	6	AX961912	AX961912 Sequence
32	40	100.0	2557	6	AX961914	AX961914 Sequence
33	40	100.0	2714	8	AB168308	AB168308 Macaca fa
34	40	100.0	5551	8	HSA012159	AJ012159 Homo sapi
35	40	100.0	7942	9	MMU012160	AJ012160 Mus muscu
36	40	100.0	121909	8	HSJ492P14	AL121977 Human DNA
37	40	100.0	167046	9	AC158516	AC158516 Mus muscu
38	40	100.0	210237	14	AC128294	AC128294 Rattus no
39	40	100.0	239076	14	AC106962	AC106962 Rattus no
40	39	97.5	1494	6	AR377516	AR377516 Sequence
41	39	97.5	158710	9	AC126539	AC126539 Mus muscu
42	39	97.5	206384	9	AC158916	AC158916 Mus muscu
43	39	97.5	208078	14	AC107258	AC107258 Rattus no
44	39	97.5	241416	14	AC114100	AC114100 Rattus no
45	38	95.0	433	6	AX385620	AX385620 Sequence
46	38	95.0	567	6	AX385964	AX385964 Sequence
47	38	95.0	3205	5	AJ851373	AJ851373 Gallus ga
48	38	95.0	40133	1	AF117351	AF117351 Zymomonas
49	38	95.0	110000	1	CR522870	Continuation (34 o
50	38	95.0	110000	1	AE008692	Continuation (19 o
51	38	95.0	110000	1	AP006627	Continuation (31 o
52	38	95.0	110000	1	AP006840	Continuation (18 o
53	38	95.0	110000	1	CP000083	Continuation (12 o
54	38	95.0	140096	8	AC092617	AC092617 Homo sapi
55	38	95.0	294800	1	SM8591789	AL591789 Sinorhizo
56	37	92.5	1320	5	BC090478	BC090478 Danio rer
57	37	92.5	110000	1	BA000016	Continuation (15 o
58	37	92.5	146190	14	AC073826	AC073826 Mus muscu
59	37	92.5	154434	5	BX649404	BX649404 Zebrafish
60	37	92.5	175371	14	AC073748	AC073748 Mus muscu
61	37	92.5	204775	9	AC151733	AC151733 Mus muscu
62	37	92.5	225010	5	BX511028	BX511028 Zebrafish
63	37	92.5	235086	14	AC097672	AC097672 Rattus no
64	37	92.5	235887	14	AC106092	AC106092 Rattus no
65	37	92.5	240016	14	AC103255	AC103255 Rattus no
66	37	92.5	243092	14	AC120099	AC120099 Rattus no
67	37	92.5	246557	14	AC148990	AC148990 Mus muscu
68	37	92.5	264260	14	AC107088	AC107088 Rattus no
69	37	92.5	302399	1	AE017164	AE017164 Prochloro
70	36	90.0	225	6	CQ650730	CQ650730 Sequence
71	36	90.0	225	6	CQ651080	CQ651080 Sequence
72	36	90.0	687	6	AR378468	AR378468 Sequence
73	36	90.0	922	2	AEVGF	L29345 Aequorea vi
74	36	90.0	922	6	AX054639	AX054639 Sequence
75	36	90.0	934	15	YSAPMP20A	J04984 C.boidinii
76	36	90.0	960	6	AR521571	AR521571 Sequence

77	36	90.0	1520	6	AR506062	AR506062 Sequence	c 150	36	90.0	238881	14	AC084103
78	36	90.0	2183	5	CR855786	CR855786 Xenopus t	c 151	36	90.0	241887	14	AC094044
79	36	90.0	2659	15	LECA88	X15258 Tomato CAB-	c 152	36	90.0	243818	14	AC134657
80	36	90.0	3222	6	CQ577362	CQ577362 Sequence	c 153	36	90.0	247243	14	AC163668
81	36	90.0	3239	2	AF145605	AF145605 Drosophila	c 154	36	90.0	283384	14	AC157810
82	36	90.0	3242	2	AY050572	AY050572 Drosophila	c 155	36	90.0	302050	14	AL935256
83	36	90.0	3352	2	BT001540	BT001540 Drosophila	c 156	36	90.0	307150	14	CNSPAX01
84	36	90.0	3491	2	AY050571	AY050571 Drosophila	c 157	36	90.0	309502	14	AC152434
85	36	90.0	3522	2	BT016131	BT016131 Drosophila	c 158	36	90.0	328451	14	AC128741
86	36	90.0	8737	6	CQ577361	CQ577361 Sequence	c 159	36	90.0	330399	2	AE003575
87	36	90.0	9220	1	AE010048	AE010048 Streptococ	c 160	36	90.0	348313	1	CR378666
88	36	90.0	11277	1	AE010008	AE010008 Streptococ	c 161	36	90.0	348505	1	BX571870
89	36	90.0	11562	1	AE009689	AE009689 Brucella	c 162	36	90.0	349800	6	AX041106
90	36	90.0	11788	1	AE006522	AE006522 Streptococ	c 163	36	90.0	349800	6	AX770909
91	36	90.0	11887	1	AE006580	AE006580 Streptococ	c 164	36	90.0	349800	6	BD204571
92	36	90.0	12126	1	AE009981	AE009981 Streptococ	c 165	36	90.0	349800	6	AE0608253
93	36	90.0	12729	1	AE007827	AE007827 Clostridi	c 166	36	90.0	349800	6	AX477784
94	36	90.0	14611	1	AE010064	AE010064 Streptococ	c 167	36	90.0	349800	6	AX066839
95	36	90.0	37497	14	AC017356	AC017356 Drosophila	c 168	36	90.0	349800	6	G26143
96	36	90.0	42050	1	AF124349	AF124349 Zymomonas	c 169	36	90.0	349800	6	G26143
97	36	90.0	50260	1	AE014155	AE014155 Streptococ	c 170	36	90.0	349800	6	BD109186
98	36	90.0	50820	1	AE014163	AE014163 Streptococ	c 171	36	90.0	349800	6	AR413633
99	36	90.0	51200	1	AE014151	AE014151 Streptococ	c 172	36	90.0	349800	6	AX970467
100	36	90.0	71409	14	AC101521	AC101521 Mus muscu	c 173	36	90.0	349800	6	BD163323
101	36	90.0	75852	8	AC125376	AC125376 Homo sapi	c 174	36	90.0	349800	6	AX121206
102	36	90.0	81346	14	AC021496	AC021496 Homo sapi	c 175	36	90.0	349800	6	C0729842
103	36	90.0	82032	14	AC120187	AC120187 Mus muscu	c 176	36	90.0	349800	6	AF151891
104	36	90.0	103694	8	AC004836	AC004836 Homo sapi	c 177	36	90.0	349800	6	AR080572
105	36	90.0	105959	15	AC137827	AC137827 Medicago	c 178	36	90.0	349800	6	BC028113
106	36	90.0	109977	14	AC091370	Continuation (6 of	c 179	36	90.0	349800	6	BD163322
107	36	90.0	110000	1	AS008692	Continuation (14 o	c 180	36	90.0	349800	6	AX121205
108	36	90.0	110000	1	AS014292	Continuation (8 of	c 181	36	90.0	349800	6	AX038988
109	36	90.0	110000	1	AS017224	Continuation (5 of	c 182	36	90.0	349800	6	AF132495
110	36	90.0	110000	1	BA000034	Continuation (23 o	c 183	36	90.0	349800	6	AR447969
111	36	90.0	110000	1	BA000034	Continuation (5 of	c 184	36	90.0	349800	6	AB169463
112	36	90.0	110000	1	BA000034	Continuation (10 o	c 185	36	90.0	349800	6	CS033983
113	36	90.0	110000	1	BA000034	Continuation (12 o	c 186	36	90.0	349800	6	CS042935
114	36	90.0	110000	1	CP000003	CP000003 Streptococ	c 187	36	90.0	349800	6	AP290514
115	36	90.0	110000	1	CP000003	Continuation (16 o	c 188	36	90.0	349800	6	AX066835
116	36	90.0	110000	1	CP000017	Continuation (12 o	c 189	36	90.0	349800	6	AR675753
117	36	90.0	110000	15	AP008214	BD430793 GENOME SE	c 190	36	90.0	349800	4	BT021631
118	36	90.0	110000	15	AP008214	Continuation (83 o	c 191	36	90.0	349800	4	BT021631
119	36	90.0	110000	15	AP008214	Continuation (84 o	c 192	36	90.0	349800	6	AB010890
120	36	90.0	118065	8	AL138783	AL138783 Human DNA	c 193	36	90.0	349800	6	CO841958
121	36	90.0	126886	14	AC145620	AC145620 Homo sapi	c 194	36	90.0	349800	8	AK123531
122	36	90.0	132958	14	AC128929	AC128929 Rattus no	c 195	36	90.0	349800	8	AB054575
123	36	90.0	142023	15	AP004705	AP004705 Oryza sat	c 196	36	90.0	349800	8	AX157302
124	36	90.0	142071	14	AC149958	AC149958 Strongylo	c 197	36	90.0	349800	8	AX933846
125	36	90.0	147223	6	CS086317	CS086317 Sequence	c 198	36	90.0	349800	8	AF164438
126	36	90.0	151882	6	CS086318	CS086318 Sequence	c 199	36	90.0	349800	6	CO363867
127	36	90.0	153657	14	AC024615	AC024615 Homo sapi	c 200	36	90.0	349800	1	AE004393
128	36	90.0	155819	8	AL358372	AL358372 Human DNA	c 201	36	90.0	349800	1	AE001900
129	36	90.0	157685	8	AC084371	AC084371 Homo sapi	c 202	36	90.0	349800	1	AE017283
130	36	90.0	161276	2	AC009749	AC009749 Drosophila	c 203	36	90.0	349800	2	CR443P11A
131	36	90.0	167908	8	AC122177	AC122177 Homo sapi	c 204	36	90.0	349800	1	MSGB1915CS
132	36	90.0	169043	14	AC018428	AC018428 Homo sapi	c 205	36	90.0	349800	6	AR345351
133	36	90.0	170563	9	AC115970	AC115970 Mus muscu	c 206	36	90.0	349800	14	T27E11
134	36	90.0	172766	8	AL513124	AL513124 Human DNA	c 207	36	90.0	349800	15	AC165222
135	36	90.0	175036	8	AC097361	AC097361 Homo sapi	c 208	36	90.0	349800	14	AC165222
136	36	90.0	188356	14	AC099550	AC099550 Homo sapi	c 209	36	90.0	349800	1	CR628336
137	36	90.0	189556	9	AC124202	AC124202 Homo sapi	c 210	36	90.0	349800	1	CR628337
138	36	90.0	190462	8	AP003057	AP003057 Homo sapi	c 211	36	90.0	349800	1	RMS591985
139	36	90.0	190488	14	AC021177	AC021177 Homo sapi	c 212	36	90.0	349800	1	AE017283
140	36	90.0	193676	14	AC018932	AC018932 Homo sapi	c 213	36	90.0	349800	1	AE017354
141	36	90.0	194068	14	AC115327	AC115327 Rattus no	c 214	36	90.0	349800	1	AP006618
142	36	90.0	198542	9	AC121912	AC121912 Mus muscu	c 215	36	90.0	349800	1	AP000022
143	36	90.0	201059	1	BSUB0017	Z99120 Bacillus su	c 216	36	90.0	349800	1	BA000022
144	36	90.0	202233	8	AC003664	AC003664 Homo sapi	c 217	36	90.0	349800	1	BA000035
145	36	90.0	223998	9	AC102534	AC102534 Mus muscu	c 218	36	90.0	349800	1	BA000036
146	36	90.0	224889	14	AC142239	AC142239 Rattus no	c 219	36	90.0	349800	1	BA000038
147	36	90.0	226531	14	AC106949	AC106949 Rattus no	c 220	36	90.0	349800	14	AC004085
148	36	90.0	235020	14	AC103022	AC103022 Rattus no	c 221	36	90.0	349800	14	AC105838
149	36	90.0	237845	14	AC112433	AC112433 Rattus no	c 222	36	90.0	349800	15	CR380953

C 223	35	87.5	110000	15	AP0080209_014	Continuation (15 o	C 296	35	87.5	249212	14	AC098919	AC098919 Rattus no
C 224	35	87.5	110941	15	AC006250	Arabidops	297	35	87.5	251122	14	AC114061	AC114061 Rattus no
C 225	35	87.5	117612	15	AC104427	Oryza sat	298	35	87.5	252151	14	AC094459	AC094459 Rattus no
C 226	35	87.5	120017	15	AC149572	Medicago	299	35	87.5	255019	14	AC162413	AC162413 Bos tauru
C 227	35	87.5	121717	8	AL445488	Human DNA	300	35	87.5	258484	14	AC125555	AC125555 Rattus no
C 228	35	87.5	133248	14	BX784399	Danio rer	301	35	87.5	260367	14	AC099427	AC099427 Rattus no
C 229	35	87.5	133476	8	AC008804	Homo sapi	C 302	35	87.5	265779	14	AC108285	AC108285 Rattus no
C 230	35	87.5	133746	14	CT009508	Medicago	303	35	87.5	267914	14	AC109865	AC109865 Rattus no
C 231	35	87.5	137496	8	AC005288	Homo sapi	304	35	87.5	277799	14	AC133319	AC133319 Rattus no
C 232	35	87.5	137667	4	AC145041	Macropus	C 305	35	87.5	300883	1	AE016809	AE016809 Vibrio vu
C 233	35	87.5	145327	8	AC114302	Homo sapi	306	35	87.5	309196	14	AC159056	AC159056 Bos tauru
C 234	35	87.5	147806	15	AC119747	Genomic s	307	35	87.5	323775	2	AE003781	AE003781 Drosophil
C 235	35	87.5	147978	8	AL136370	Human DNA	C 308	35	87.5	349459	1	BX927151	BX927151 Corynebac
C 236	35	87.5	149005	15	AC146785	Medicago	C 309	35	87.5	349483	1	BX571859	BX571859 Photorhab
C 237	35	87.5	153225	14	AC013581	Homo sapi	C 310	35	87.5	349860	6	AX573241	AX573241 Sequence
C 238	35	87.5	153297	14	AC027558	Homo sapi	C 311	35	87.5	349860	6	AX127146	AX127146 Sequence
C 239	35	87.5	153538	8	AC0933279	Homo sapi	C 312	34	85.0	201	10	BV182913	BV182913 sqm13415
C 240	35	87.5	154227	14	CR628323	Danio rer	C 313	34	85.0	213	6	CB748469	CB748469 Sequence
C 241	35	87.5	156763	8	AC005901	Homo sapi	C 314	34	85.0	390	5	AB101390	AB101390 Haplochro
C 242	35	87.5	159468	14	AC015280	Drosophil	C 315	34	85.0	421	10	HSPA3289	Z94308 H. sapiens f
C 243	35	87.5	161938	8	CNS0180Y	Human chr	C 316	34	85.0	511	10	BV102769	BV102769 MARC 2065
C 244	35	87.5	165423	14	AC008811	Homo sapi	C 317	34	85.0	608	10	BV418280	BV418280 S229P6362
C 245	35	87.5	165716	5	CR376858	Zebrafish	C 318	34	85.0	644	10	BV038609	BV038609 S212P6032
C 246	35	87.5	166054	9	AC134539	Mus muscu	C 319	34	85.0	750	2	AK174110	AK174110 Ciona int
C 247	35	87.5	167042	8	AC117527	Homo sapi	C 320	34	85.0	904	6	BD069416	BD069416 Mammalian
C 248	35	87.5	167195	8	AC009044	Homo sapi	C 321	34	85.0	904	6	AR228066	AR228066 Sequence
C 249	35	87.5	168839	8	AC022826	Homo sapi	C 322	34	85.0	929	15	YSAFMP20B	J04985 Yeast (C.bo
C 250	35	87.5	169358	14	CR318602	Danio rer	C 323	34	85.0	1079	5	AF092937	AF092937 Acipenser
C 251	35	87.5	170327	2	AC006467	Drosophil	C 324	34	85.0	1055	5	D88740	D88740 Acipenser t
C 252	35	87.5	170520	14	AL358172	Homo sapi	C 325	34	85.0	1140	8	AF271413	AF271413 Perodicti
C 253	35	87.5	172143	8	AC113949	Homo sapi	C 326	34	85.0	1140	8	AY441473	AY441473 Perodicti
C 254	35	87.5	174039	8	AC138853	Homo sapi	C 327	34	85.0	1140	8	AY687891	AY687891 Nycticebu
C 255	35	87.5	175320	8	AC009283	Homo sapi	C 328	34	85.0	1140	8	AY687892	AY687892 Nycticebu
C 256	35	87.5	176277	14	AC139815	Homo sapi	C 329	34	85.0	1140	8	AY687893	AY687893 Nycticebu
C 257	35	87.5	177012	14	AC020989	Homo sapi	C 330	34	85.0	1140	8	AY687894	AY687894 Nycticebu
C 258	35	87.5	179099	2	AC150553	Drosophil	C 331	34	85.0	1140	8	AY687900	AY687900 Nycticebu
C 259	35	87.5	179399	5	AL929338	Zebrafish	C 332	34	85.0	1476	2	AY363117	AY363117 Caenorhab
C 260	35	87.5	181217	14	AC120674	Rattus no	C 333	34	85.0	1550	2	AF183391	AF183391 Caenorhab
C 261	35	87.5	182588	9	AC137107	Mus muscu	C 334	34	85.0	1550	2	AF183391	AF183391 Caenorhab
C 262	35	87.5	182736	5	BX511146	Zebrafish	C 335	34	85.0	2247	6	AC056396	AC056396 Sequence
C 263	35	87.5	183369	9	AL591205	Mouse DNA	C 336	34	85.0	2271	8	BC064614	BC064614 Homo sapi
C 264	35	87.5	186413	9	AC121769	Mus muscu	C 337	34	85.0	2283	15	AV124871	AV124871 Arabidops
C 265	35	87.5	187409	9	AC151572	Mus muscu	C 338	34	85.0	2315	7	AB182649	AB182649 Bacteriop
C 266	35	87.5	188037	5	BX545913	Mus muscu	C 339	34	85.0	2408	15	AB055180	AB055180 Candida b
C 267	35	87.5	189425	14	AC142182	Rattus no	C 340	34	85.0	2482	15	AY057700	AY057700 Arabidops
C 268	35	87.5	190050	1	AL646080	Rattus no	C 341	34	85.0	2520	5	BC044429	BC044429 Danio rer
C 269	35	87.5	190676	14	CR847904	Danio rer	C 342	34	85.0	2533	4	PIGASPAT	PIGASPAT Pig mitoch
C 270	35	87.5	190684	14	AC139798	Homo sapi	C 343	34	85.0	2597	1	AF109471	AF109471 Xanthomon
C 271	35	87.5	193666	14	AC140124	Homo sapi	C 344	34	85.0	3190	15	AK073114	AK073114 Oryza sat
C 272	35	87.5	193879	14	AC140510	Homo sapi	C 345	34	85.0	3208	15	AK100158	AK100158 Oryza sat
C 273	35	87.5	198157	14	AC125888	Rattus no	C 346	34	85.0	3382	1	LFU48430	U48430 Lactobacill
C 274	35	87.5	199749	15	ATC8RIV68	Arabidops	C 347	34	85.0	3577	15	AK071459	AK071459 Oryza sat
C 275	35	87.5	200001	15	ATC8RIV67	Arabidops	C 348	34	85.0	4467	6	AR043598	AR043598 Sequence
C 276	35	87.5	200380	8	AC022898	Homo sapi	C 349	34	85.0	4467	6	AR071336	AR071336 Sequence
C 277	35	87.5	201542	9	AL732483	Mouse DNA	C 350	34	85.0	7858	1	U35629	U35629 Laccococcu
C 278	35	87.5	203786	14	AC139809	Homo sapi	C 351	34	85.0	10312	1	AE015007	AE015007 Streptoco
C 279	35	87.5	208532	14	AC128540	Rattus no	C 352	34	85.0	11153	1	AE015747	AE015747 Shewanell
C 280	35	87.5	213029	9	AC154688	Mus muscu	C 353	34	85.0	11427	1	AE010884	AE010884 Methanosa
C 281	35	87.5	219498	14	AC153209	Bos tauru	C 354	34	85.0	12082	1	AE013606	AE013606 Versinia
C 282	35	87.5	220663	14	AC094574	Rattus no	C 355	34	85.0	12083	1	AE002012	AE002012 Deinococc
C 283	35	87.5	221145	14	AC112859	Rattus no	C 356	34	85.0	12371	9	MMMDMPK	Z38015 M. musculus
C 284	35	87.5	223728	14	AC135443	Rattus no	C 357	34	85.0	15267	2	AY191994	AY191994 Tricholep
C 285	35	87.5	226299	14	AC113230	Rattus no	C 358	34	85.0	18209	1	AE005129	AE005129 Halobacte
C 286	35	87.5	227401	14	AC105613	Rattus no	C 359	34	85.0	31235	9	BX571789	BX571789 Mouse DNA
C 287	35	87.5	227816	9	AL683884	Mouse DNA	C 360	34	85.0	36720	14	AC165829	AC165829 Bos tauru
C 288	35	87.5	228126	14	AC153052	Bos tauru	C 361	34	85.0	37900	1	D86417	D86417 Bacillus su
C 289	35	87.5	228126	14	AC153052	Bos tauru	C 362	34	85.0	38145	2	U64847	U64847 Caenorhabdi
C 290	35	87.5	231785	14	AC094057	Rattus no	C 363	34	85.0	38450	8	BS000570	BS000570 Pan trogl
C 291	35	87.5	235933	14	AC097843	Rattus no	C 364	34	85.0	45292	8	AC114957	AC114957 Homo sapi
C 292	35	87.5	237893	14	AC094418	Rattus no	C 365	34	85.0	49999	6	AX015902	AX015902 Sequence
C 293	35	87.5	238045	14	AC096146	Rattus no	C 366	34	85.0	49999	6	AX015908	AX015908 Sequence
C 294	35	87.5	238097	14	AC127410	Rattus no	C 367	34	85.0	52684	15	AF111709	AF111709 Oryza sat
C 295	35	87.5	247017	14	AC110384	Rattus no	C 368	34	85.0	60805	14	AC109323	AC109323 Homo sapi

c 369	34	85.0	61429	14	AC104239	AC104239 Homo sapi	442	34	85.0	113672	8	AC010455	AC010455 Homo sapi
370	34	85.0	62050	14	AC124823	AC124823 Mus muscu	443	34	85.0	115005	15	AC134242	AC134242 Medicago
c 371	34	85.0	64707	2	AC115607	AC115607 Dictyoste	c 444	34	85.0	117273	8	AL451138	AL451138 Human DNA
372	34	85.0	67159	14	AC103698	AC103698 Homo sapi	445	34	85.0	117580	14	AP006455	AP006455 Oryza sat
c 373	34	85.0	67159	14	AC103698	AC103698 Homo sapi	446	34	85.0	117827	14	AC152366	AC152366 Felis cat
374	34	85.0	67896	14	AC101995	AC101995 Mus muscu	c 447	34	85.0	118185	9	AL627235	AL627235 Mouse DNA
375	34	85.0	67922	8	AC109324	AC109324 Homo sapi	448	34	85.0	118699	14	AY772735	AY772735 Trifolium
c 376	34	85.0	73073	14	AC166762	AC166762 Bos tauru	449	34	85.0	120630	8	HS28F12	AC01657 Homo sapi
377	34	85.0	73940	8	AC1233905	AC1233905 Homo sapi	c 450	34	85.0	120994	8	AC008697	AC008697 Homo sapi
c 378	34	85.0	75969	14	AC151697	AC151697 Gallus ga	451	34	85.0	122504	8	AC106757	AC106757 Homo sapi
379	34	85.0	78376	8	AL359692	AL359692 Human DNA	c 452	34	85.0	125794	14	AC160616	AC160616 Loxodonta
c 380	34	85.0	79034	14	AL731895_3	Continuation (4 of	453	34	85.0	125799	8	AC004682	AC004682 Homo sapi
c 381	34	85.0	80115	14	AP008123	AP008123 Lotus cor	454	34	85.0	125799	8	AL935327	AL935327 Mouse DNA
c 382	34	85.0	82166	8	AC006237	AC006237 Homo sapi	c 455	34	85.0	127776	14	AC133780	AC133780 Medicago
c 383	34	85.0	83177	3	AY458647	AY458647 Unculture	456	34	85.0	128520	15	AP004123	AP004123 Oryza sat
c 384	34	85.0	84081	5	AL672016	AL672016 Zebrafish	457	34	85.0	128544	8	AF238378	AF238378 Homo sapi
c 385	34	85.0	84182	8	AC110763	AC110763 Homo sapi	458	34	85.0	129206	14	AL121940	AL121940 Homo sapi
c 386	34	85.0	85427	14	HSDJ470K3	AL109655 Homo sapi	459	34	85.0	131271	14	AC015927	AC015927 Homo sapi
c 387	34	85.0	88068	15	AC002337	AC002337 Arabidops	c 460	34	85.0	132983	14	AC128205	AC128205 Rattus no
c 388	34	85.0	89566	9	CR931957	CR931957 Mouse DNA	c 461	34	85.0	133362	8	AL139230	AL139230 Human DNA
c 389	34	85.0	91148	14	AC119727_3	Continuation (4 of	462	34	85.0	134372	9	AL645805	AL645805 Mouse DNA
c 390	34	85.0	94175	14	AC136849	AC136849 Rattus no	463	34	85.0	134415	8	AC087207	AC087207 Homo sapi
c 391	34	85.0	95118	15	AC146746	AC146746 Medicago	464	34	85.0	136862	14	AC146853	AC146853 Medicago
c 392	34	85.0	97572	15	AP004155	AP004155 Oryza sat	465	34	85.0	136989	9	AC116710	AC116710 Mus muscu
c 393	34	85.0	98328	8	AC113372	AC113372 Homo sapi	466	34	85.0	137580	14	AC160243	AC160243 Mus muscu
c 394	34	85.0	102842	15	AF474071	AF474071 Hordeum v	467	34	85.0	137764	14	AC163962	AC163962 Loxodonta
c 395	34	85.0	104475	8	AC011384	AC011384 Homo sapi	468	34	85.0	138093	8	AL606467	AL606467 Human DNA
c 396	34	85.0	108039	8	AC024565	AC024565 Homo sapi	c 469	34	85.0	140553	15	AC091666	AC091666 Oryza sat
c 397	34	85.0	108511	8	AC011418	AC011418 Homo sapi	c 470	34	85.0	140718	8	AL583839	AL583839 Human DNA
c 398	34	85.0	108545	14	AC148376	AC148376 Sorex ara	c 471	34	85.0	141415	8	AC087214	AC087214 Papio anu
c 399	34	85.0	108976	8	AC008604	AC008604 Homo sapi	472	34	85.0	141580	15	OSJN00278	AP005413 Oryza sat
c 400	34	85.0	110000	1	CP000099_32	Continuation (33 o	473	34	85.0	142702	15	AP005413	AP005413 Oryza sat
c 401	34	85.0	110000	1	CR555306_22	Continuation (23 o	c 474	34	85.0	143153	14	AC025027	AC025027 Homo sapi
c 402	34	85.0	110000	1	CR931997_02	Continuation (3 of	c 475	34	85.0	143599	14	AC012947	AC012947 Drosophila
c 403	34	85.0	110000	1	CR931997_19	Continuation (20 o	c 476	34	85.0	144430	8	AC132803	AC132803 Homo sapi
c 404	34	85.0	110000	1	AE016827_07	Continuation (8 of	c 477	34	85.0	144709	14	AC120084	AC120084 Rattus no
c 405	34	85.0	110000	1	AE017180_30	Continuation (31 o	478	34	85.0	144947	14	AC162909	AC162909 Mus muscu
c 406	34	85.0	110000	1	AE017340_00	Continuation (4 of	479	34	85.0	147817	15	AP003267	AP003267 Oryza sat
c 407	34	85.0	110000	1	AF006716_03	Continuation (4 of	c 480	34	85.0	147913	14	AC026971	AC026971 Homo sapi
c 408	34	85.0	110000	1	BA000001_00	BA000001 Pyrococu	c 481	34	85.0	148752	15	AP005586	AP005586 Oryza sat
c 409	34	85.0	110000	1	BA000011_02	Continuation (3 of	c 482	34	85.0	149505	8	AC126927	AC126927 Felis cat
c 410	34	85.0	110000	1	BA000028_04	Continuation (5 of	483	34	85.0	152829	8	AC078793	AC078793 Homo sapi
c 411	34	85.0	110000	1	BA000038_06	Continuation (7 of	484	34	85.0	155976	14	AC055113	AC055113 Homo sapi
c 412	34	85.0	110000	1	BA000078_03	Continuation (4 of	c 485	34	85.0	156842	14	AC069318	AC069318 Homo sapi
c 413	34	85.0	110000	1	BA000098_00	Continuation (2 of	c 486	34	85.0	156895	14	AC120107	AC120107 Rattus no
c 414	34	85.0	110000	1	BA000098_01	Continuation (7 of	c 487	34	85.0	157035	14	AC160531	AC160531 Mus muscu
c 415	34	85.0	110000	1	CP000001_06	Continuation (20 o	c 488	34	85.0	157520	14	AC114317	AC114317 Homo sapi
c 416	34	85.0	110000	1	CP000009_19	Continuation (6 of	c 489	34	85.0	157772	8	AC006212	AC006212 Homo sapi
c 417	34	85.0	110000	1	CP000027_05	Continuation (7 of	c 490	34	85.0	158209	8	AC105214	AC105214 Homo sapi
c 418	34	85.0	110000	1	CP000027_06	Continuation (7 of	c 491	34	85.0	159046	14	AC092426	AC092426 Homo sapi
c 419	34	85.0	110000	1	CP000031_21	Continuation (22 o	c 492	34	85.0	159617	8	AC022018	AC022018 Homo sapi
c 420	34	85.0	110000	1	CP000076_67	Continuation (68 o	493	34	85.0	160246	14	AC009221	AC009221 Homo sapi
c 421	34	85.0	110000	14	AC091229_09	Continuation (11 o	c 494	34	85.0	160403	9	AC118940	AC118940 Mus muscu
c 422	34	85.0	110000	14	AC091229_10	Continuation (2 of	c 495	34	85.0	160436	8	AC146454	AC146454 Pan trogl
c 423	34	85.0	110000	14	AC098564_1	Continuation (3 of	c 496	34	85.0	160850	8	AC130184	AC130184 Macaca mu
c 424	34	85.0	110000	14	AC128152_2	Continuation (4 of	c 497	34	85.0	161475	8	HS164C20	AL009028 Human DNA
c 425	34	85.0	110000	14	AC149098_3	Continuation (4 of	c 498	34	85.0	161831	9	AC156558	AC156558 Mus muscu
c 426	34	85.0	110000	14	CR542404_2	AL732359 Homo sapi	499	34	85.0	162129	9	AC154488	AC154488 Mus muscu
c 427	34	85.0	110000	15	AP008214_094	Continuation (3 of	500	34	85.0	162320	14	AC069225	AC069225 Homo sapi
c 428	34	85.0	110000	15	AP008214_082	Continuation (95 o	c 501	34	85.0	162776	15	AP003237	AP003237 Oryza sat
c 429	34	85.0	110000	15	AP008216_099	Continuation (83 o	c 502	34	85.0	163778	9	AC123682	AC123682 Mus muscu
c 430	34	85.0	110000	15	AP008216_131	Continuation (100	c 503	34	85.0	163197	5	EX005416	EX005416 Zebrafish
c 431	34	85.0	110000	15	AP008216_131	Continuation (132	c 504	34	85.0	163944	14	AC156770	AL845513 Zebrafish
c 432	34	85.0	110000	15	CR380949_3	Continuation (4 of	c 505	34	85.0	164117	5	AL845513	AL845513 Zebrafish
c 433	34	85.0	110000	15	AP008213_21	Continuation (22 o	506	34	85.0	164460	5	AC144822	AC144822 Danio rer
c 434	34	85.0	110000	15	AP008207_213	Continuation (214	c 507	34	85.0	165900	14	AP003942	AP003942 Oryza sat
c 435	34	85.0	110000	15	AP008207_314	Continuation (315	c 508	34	85.0	166417	9	AC127685	AC127685 Mus muscu
c 436	34	85.0	110000	15	AP008208_114	Continuation (115	c 509	34	85.0	166506	2	AY449462	AY449462 Oikopleur
c 437	34	85.0	110000	15	AP008210_026	Continuation (27 o	c 510	34	85.0	166619	14	AC142132	AC142132 Rattus no
c 438	34	85.0	110000	15	AP008210_112	Continuation (113	c 511	34	85.0	167384	9	AC138640	AC138640 Mus muscu
c 439	34	85.0	110000	15	AP008212_108	Continuation (109	c 512	34	85.0	168049	14	AC022830	AC022830 Homo sapi
c 440	34	85.0	112084	8	AC104648	AC104648 Homo sapi	c 513	34	85.0	168601	14	AC123445	AC123445 Rattus no
c 441	34	85.0	112124	8	AL137184	AL137184 Human DNA	c 514	34	85.0	168963	14	AC009705	AC009705 Homo sapi

515	34	85.0	169270	15	AC083943	Genomic s	AC083943	Genomic s	34	85.0	207992	14	AC155613	AC155613	Zea mays
516	34	85.0	170059	14	AC128291	Rattus no	AC128291	Rattus no	34	85.0	208598	14	AC160508	AC160508	Bos tauru
C 517	34	85.0	170996	14	AC012686	Homo sapi	AC012686	Homo sapi	34	85.0	213708	14	AC110888	AC110888	Mus muscu
C 518	34	85.0	171818	14	AC102862	Mus muscu	AC102862	Mus muscu	34	85.0	214824	14	AC106466	AC106466	Rattus no
C 519	34	85.0	171944	14	AC138724	Cercopith	AC138724	Cercopith	592	85.0	216062	14	AC103262	AC103262	Rattus no
C 520	34	85.0	172022	14	AC145440	Bos tauru	AC145440	Bos tauru	C 593	85.0	217476	2	AE003685	AE003685	Drosophil
C 521	34	85.0	172490	9	AC138595	Mus muscu	AC138595	Mus muscu	C 594	85.0	217976	14	AC106637	AC106637	Rattus no
C 522	34	85.0	174450	8	AC159110	Pan trogl	AC159110	Pan trogl	C 595	85.0	218778	14	AC137406	AC137406	Rattus no
C 523	34	85.0	174648	4	AC150651	Bos tauru	AC150651	Bos tauru	C 596	85.0	218877	14	AC165303	AC165303	Mus muscu
C 524	34	85.0	175593	15	OSJUN00252		AL731603	Oryza sat	C 597	85.0	220621	5	CR628387	CR628387	Zebrafish
C 525	34	85.0	175820	8	ALJ162378	Human DNA	ALJ162378	Human DNA	C 598	85.0	220747	14	AC110120	AC110120	Rattus no
C 526	34	85.0	176018	14	AC157755	Pan trogl	AC157755	Pan trogl	C 599	85.0	220793	14	AC130046	AC130046	Rattus no
C 527	34	85.0	176050	14	AC131279	Homo sapi	AC131279	Homo sapi	C 600	85.0	220932	9	AC145199	AC145199	Mus muscu
C 528	34	85.0	176247	14	AC101805	Mus muscu	AC101805	Mus muscu	601	85.0	223031	14	AC113837	AC113837	Rattus no
C 529	34	85.0	177750	14	AC040947	Homo sapi	AC040947	Homo sapi	602	85.0	223455	14	AC160513	AC160513	Rattus no
C 530	34	85.0	178153	9	AL954346		AL954346	Mouse DNA	603	85.0	223670	1	CR555308	CR555308	Azarcus
C 531	34	85.0	179680	8	AC016138	Homo sapi	AC016138	Homo sapi	604	85.0	227192	14	AC163718	AC163718	Pan trogl
C 532	34	85.0	179756	14	AC074039	Mus muscu	AC074039	Mus muscu	C 605	85.0	228283	9	AF312994	AF312994	Mus muscu
C 533	34	85.0	179966	14	AC023247		AC023247	Homo sapi	C 606	85.0	228375	14	AC114460	AC114460	Rattus no
C 534	34	85.0	180227	5	CR769781		CR769781	Zebrafish	607	85.0	230056	14	AC134124	AC134124	Rattus no
C 535	34	85.0	180679	8	AC091900	Homo sapi	AC091900	Homo sapi	C 608	85.0	230299	14	AC098763	AC098763	Rattus no
C 536	34	85.0	180838	8	AC159109	Pan trogl	AC159109	Pan trogl	C 609	85.0	231316	14	AC110446	AC110446	Rattus no
C 537	34	85.0	181197	9	AC160467	Mus muscu	AC160467	Mus muscu	610	85.0	231660	14	AC108236	AC108236	Rattus no
C 538	34	85.0	181464	14	AC154167	Rhinoloph	AC154167	Rhinoloph	C 611	85.0	231817	14	AC137474	AC137474	Rattus no
C 539	34	85.0	181564	14	CR855880	Danio rer	CR855880	Danio rer	612	85.0	233094	14	AC134131	AC134131	Rattus no
C 540	34	85.0	181700	8	AC009965	Homo sapi	AC009965	Homo sapi	613	85.0	233399	14	AC095527	AC095527	Rattus no
C 541	34	85.0	182161	9	AC158912	Mus muscu	AC158912	Mus muscu	C 614	85.0	233553	14	AC127214	AC127214	Rattus no
C 542	34	85.0	182846	14	AC068223	Homo sapi	AC068223	Homo sapi	615	85.0	233825	14	AC129679	AC129679	Rattus no
C 543	34	85.0	183320	14	AC109867	Rattus no	AC109867	Rattus no	C 616	85.0	235777	14	AC096346	AC096346	Rattus no
C 544	34	85.0	183438	8	AC007738	Homo sapi	AC007738	Homo sapi	C 617	85.0	235804	14	AC127084	AC127084	Rattus no
C 545	34	85.0	183577	14	AC001164	Homo sapi	AC001164	Homo sapi	618	85.0	236466	14	AC137243	AC137243	Rattus no
C 546	34	85.0	183632	14	AC157470	Rhinoloph	AC157470	Rhinoloph	619	85.0	237954	14	AC128532	AC128532	Rattus no
C 547	34	85.0	183814	14	AC155975	Xenopus t	AC155975	Xenopus t	C 620	85.0	238462	14	AC160064	AC160064	Rattus no
C 548	34	85.0	185030	14	AC149009	Papio anu	AC149009	Papio anu	C 621	85.0	238731	14	AC108666	AC108666	Rattus no
C 549	34	85.0	185073	14	AC160225	Loxodonta	AC160225	Loxodonta	622	85.0	240503	14	AC111465	AC111465	Rattus no
C 550	34	85.0	185770	14	AC147947	Macropus	AC147947	Macropus	623	85.0	240934	14	AC096409	AC096409	Rattus no
C 551	34	85.0	187303	8	AC069226	Homo sapi	AC069226	Homo sapi	C 624	85.0	241644	14	AC094176	AC094176	Rattus no
C 552	34	85.0	188681	14	AC166878	Oryctolag	AC166878	Oryctolag	C 625	85.0	241765	14	AC132557	AC132557	Rattus no
C 553	34	85.0	189479	8	AC027309	Homo sapi	AC027309	Homo sapi	626	85.0	242460	14	AC095535	AC095535	Rattus no
C 554	34	85.0	190015	9	AC102191	Mus muscu	AC102191	Mus muscu	627	85.0	243259	14	AC105355	AC105355	Rattus no
C 555	34	85.0	190151	14	AC149113	Papio anu	AC149113	Papio anu	628	85.0	243812	14	AC118510	AC118510	Rattus no
C 556	34	85.0	190580	8	AC010184	Homo sapi	AC010184	Homo sapi	C 629	85.0	244140	14	AC128754	AC128754	Rattus no
C 557	34	85.0	191020	14	AC146845	Otolemur	AC146845	Otolemur	630	85.0	244409	14	AC163131	AC163131	Bos tauru
C 558	34	85.0	191790	9	AC121773	Mus muscu	AC121773	Mus muscu	631	85.0	246096	14	AC126715	AC126715	Rattus no
C 559	34	85.0	192055	2	AC009183	Drosophil	AC009183	Drosophil	C 632	85.0	246134	14	AC022082	AC022082	Homo sapi
C 560	34	85.0	192187	9	AC113441	Mus muscu	AC113441	Mus muscu	633	85.0	246337	14	AC106464	AC106464	Rattus no
C 561	34	85.0	193960	8	AC022217	Homo sapi	AC022217	Homo sapi	634	85.0	246826	14	AC116225	AC116225	Rattus no
C 562	34	85.0	194115	5	CR391911	Zebrafish	CR391911	Zebrafish	635	85.0	246975	14	AC129171	AC129171	Rattus no
C 563	34	85.0	194120	5	AL953844		AL953844	Zebrafish	636	85.0	247037	14	AC094028	AC094028	Rattus no
C 564	34	85.0	194641	9	AC124107	Mus muscu	AC124107	Mus muscu	C 637	85.0	247108	14	AC123333	AC123333	Rattus no
C 565	34	85.0	195029	14	AC007902	Homo sapi	AC007902	Homo sapi	C 638	85.0	249251	9	AC134585	AC134585	Mus muscu
C 566	34	85.0	195987	9	AC122537	Mus muscu	AC122537	Mus muscu	639	85.0	249658	14	AC133658	AC133658	Rattus no
C 567	34	85.0	197398	9	AL732360	Mouse DNA	AL732360	Mouse DNA	C 640	85.0	250399	14	AC126813	AC126813	Rattus no
C 568	34	85.0	197409	1	BSUB0005	Bacillus su	299108	Bacillus su	C 641	85.0	251402	14	AC112002	AC112002	Rattus no
C 569	34	85.0	197429	14	AC023666	Homo sapi	AC023666	Homo sapi	642	85.0	252421	14	AC106361	AC106361	Rattus no
C 570	34	85.0	197778	14	AC165380	Colobus g	AC165380	Colobus g	C 643	85.0	254977	14	AC117353	AC117353	Rattus no
C 571	34	85.0	198712	14	AC160566	Pan trogl	AC160566	Pan trogl	C 644	85.0	255998	14	AC125769	AC125769	Rattus no
C 572	34	85.0	200368	9	AL671895	Mouse DNA	AL671895	Mouse DNA	645	85.0	258317	14	AC129732	AC129732	Rattus no
C 573	34	85.0	200493	8	AP005270	Homo sapi	AP005270	Homo sapi	646	85.0	258635	14	AC106208	AC106208	Rattus no
C 574	34	85.0	200721	14	AC119742	Homo sapi	AC119742	Homo sapi	C 647	85.0	260924	14	AC095740	AC095740	Rattus no
C 575	34	85.0	202214	14	AC135818	Rattus no	AC135818	Rattus no	C 648	85.0	261162	1	AB017257	AB017257	Wolbachia
C 576	34	85.0	202727	14	AC140003	Rattus no	AC140003	Rattus no	649	85.0	261498	14	AC073823	AC073823	Mus muscu
C 577	34	85.0	203278	8	AC010482	Homo sapi	AC010482	Homo sapi	650	85.0	263468	14	AC106322	AC106322	Rattus no
C 578	34	85.0	203348	8	AC104376	Homo sapi	AC104376	Homo sapi	651	85.0	264819	14	AC098529	AC098529	Rattus no
C 579	34	85.0	203568	8	AC008693	Homo sapi	AC008693	Homo sapi	652	85.0	266916	14	CR388402	CR388402	Danio rer
C 580	34	85.0	205040	8	AC015849	Homo sapi	AC015849	Homo sapi	653	85.0	267531	14	AC127708	AC127708	Rattus no
C 581	34	85.0	205472	14	AC161385	Pan trogl	AC161385	Pan trogl	C 654	85.0	269251	14	AC097289	AC097289	Rattus no
C 582	34	85.0	206094	9	AC119861	Mus muscu	AC119861	Mus muscu	655	85.0	272677	14	AC097289	AC097289	Rattus no
C 583	34	85.0	206240	9	AC132453	Mus muscu	AC132453	Mus muscu	656	85.0	272686	14	AC130874	AC130874	Rattus no
C 584	34	85.0	206470	14	AC135811	Rattus no	AC135811	Rattus no	657	85.0	274492	14	AC159688	AC159688	Bos tauru
C 585	34	85.0	206891	14	AC122575	Rattus no	AC122575	Rattus no	C 658	85.0	288050	1	AJ414141	AJ414141	Yersinia
C 586	34	85.0	206898	9	AC152949	Mus muscu	AC152949	Mus muscu	659	85.0	289080	14	AC111582	AC111582	Rattus no
C 587	34	85.0	207313	14	AC121620	Rattus no	AC121620	Rattus no	C 660	85.0	290029	1	AB017127	AB017127	Yersinia

C 661	34	85.0	292171	14	AC091244	AC091244 Rattus no	734	33	82.5	1514	4	PTGFLTP	L78843 Sus scrofa
C 662	34	85.0	295854	8	AF228730	Homo sapi	735	33	82.5	1528	1	AX134440	AX134440 Candidatu
C 663	34	85.0	300029	15	AE017080	Oryza sat	736	33	82.5	1728	2	AX112275	AX112275 Ciona int
C 664	34	85.0	301936	1	AE016808	Vibrio vu	737	33	82.5	1874	13	AY821804	AY821804 Caprine h
C 665	34	85.0	308147	1	AE016915	Chromobac	738	33	82.5	1902	6	AR107058	AR107058 Sequence
C 666	34	85.0	312988	14	CR974592	Mus muscu	739	33	82.5	1902	6	AR107059	AR107059 Sequence
C 667	34	85.0	324367	15	AE017091	Oryza sat	740	33	82.5	1902	6	BD082244	BD082244 Bacillus
C 668	34	85.0	326249	14	AC034210	Homo sapi	741	33	82.5	1902	6	BD082245	BD082245 Bacillus
C 669	34	85.0	326449	14	AC115567	Rattus no	742	33	82.5	2000	6	AX594862	AX594862 Sequence
C 670	34	85.0	326700	14	AC113462	Rattus no	743	33	82.5	2000	6	AX818748	AX818748 Sequence
C 671	34	85.0	328867	14	AC097130	Rattus no	744	33	82.5	2000	6	AX829778	AX829778 Sequence
C 672	34	85.0	333622	14	AC111475	Rattus no	745	33	82.5	2007	13	CHV10C	Z49225 Caprine her
C 673	34	85.0	349497	1	BX640440	Bordetell	746	33	82.5	2084	9	AF326772	AF326772 Rattus no
C 674	33	82.5	213	6	AX396097	Sequence	747	33	82.5	2176	9	BC079379	BC079379 Rattus no
C 675	33	82.5	273	6	AX396024	Sequence	748	33	82.5	2234	6	AR227602	AR227602 Sequence
C 676	33	82.5	294	6	AR348638	Sequence	749	33	82.5	2235	9	BC062199	BC062199 Mus muscu
C 677	33	82.5	315	3	AY309218	Unculture	750	33	82.5	2236	9	BC036178	BC036178 Mus muscu
C 678	33	82.5	401	10	BV189117	sgmml6242	751	33	82.5	2395	1	AF123492	AF123492 Pseudomon
C 679	33	82.5	401	10	BV189351	sgmml6334	752	33	82.5	2410	9	BC088860	BC088860 Rattus no
C 680	33	82.5	447	2	AF092683	Neochlami	753	33	82.5	2430	9	BC036179	BC036179 Mus muscu
C 681	33	82.5	474	11	AY774381	Synthetic	754	33	82.5	2578	6	BC652158	BC652158 Sequence
C 682	33	82.5	478	10	G95845	S210P6025RD	755	33	82.5	2783	9	BC039183	BC039183 Mus muscu
C 683	33	82.5	496	10	G88118	S208P6566FA	756	33	82.5	2819	6	C0604148	C0604148 Sequence
C 684	33	82.5	531	6	AX396096	Sequence	757	33	82.5	2823	15	AX101115	AX101115 Sequence
C 685	33	82.5	561	10	BV341359	S330P6555	758	33	82.5	3009	6	C0609639	C0609639 Sequence
C 686	33	82.5	584	10	G80680	S208P6453FA	759	33	82.5	3010	1	AY486146	AY486146 Lactobacil
C 687	33	82.5	615	6	C0736619	Sequence	760	33	82.5	3111	2	AGA535207	AGA535207 Anopheles
C 688	33	82.5	618	10	BV376239	S231P6394	761	33	82.5	3121	1	AF200816	AF200816 Bacillus
C 689	33	82.5	690	11	AY658047	Synthetic	762	33	82.5	3755	2	AY151155	AY151155 Loligo pe
C 690	33	82.5	692	1	AF435437	Enterococ	763	33	82.5	3852	15	SCU31331	SCU31331 Saccharomyc
C 691	33	82.5	692	1	AF435438	Enterococ	764	33	82.5	4618	9	AK220531	AK220531 Mus muscu
C 692	33	82.5	692	1	AF435439	Enterococ	765	33	82.5	5293	6	C0609638	C0609638 Sequence
C 693	33	82.5	692	1	AF435440	Enterococ	766	33	82.5	5580	1	CCCEURBCDE	X88849 C.coli ceuB
C 694	33	82.5	692	1	AF435441	Enterococ	767	33	82.5	5598	6	BD193525	BD193525 Enterococ
C 695	33	82.5	692	1	AF435442	Enterococ	768	33	82.5	7480	8	AB111888	AB111888 Homo sapi
C 696	33	82.5	705	15	MGR299236	AF435442 Enterococ	769	33	82.5	7574	6	BD193558	BD193558 Enterococ
C 697	33	82.5	735	1	BSTRNA1	X00899 Bacillus su	770	33	82.5	8442	1	AE074151	AE074151 Vibrio vi
C 698	33	82.5	744	6	AR477818	Sequence	771	33	82.5	8873	15	SCYBL088C	Z35849 S.cerevisia
C 699	33	82.5	744	6	AX066907	Sequence	772	33	82.5	9424	1	HUI17295	UI7295 Haemophilus
C 700	33	82.5	750	6	AR550949	Sequence	773	33	82.5	9462	8	AL358412	AL358412 Human DNA
C 701	33	82.5	798	6	AR552159	Sequence	774	33	82.5	9836	1	U93690	U93690 Mycoplasma
C 702	33	82.5	819	6	CQ604149	Sequence	775	33	82.5	10016	1	AE004987	AE004987 Halobacte
C 703	33	82.5	853	1	AY236070	Ehrlichia	776	33	82.5	10081	1	AE004512	AE004512 Pseudomon
C 704	33	82.5	880	6	BD261390	Chlamydia	777	33	82.5	10263	1	AE013924	AE013924 Yersinia
C 705	33	82.5	882	6	AR464502	Sequence	778	33	82.5	10423	1	U32798	U32798 Haemophilus
C 706	33	82.5	897	6	AR558503	Sequence	779	33	82.5	10432	1	AE006885	AE006885 Sulfolobu
C 707	33	82.5	993	1	AY236059	Ehrlichia	780	33	82.5	10459	1	AE004717	AE004717 Pseudomon
C 708	33	82.5	993	1	AY236060	Ehrlichia	781	33	82.5	10500	1	NME391262	AE004717 Pseudomon
C 709	33	82.5	993	1	AY236061	Ehrlichia	782	33	82.5	10509	6	CQ872934	CQ872934 Sequence
C 710	33	82.5	993	1	AY236062	Ehrlichia	783	33	82.5	10676	1	AE008293	AE008293 Agrobacte
C 711	33	82.5	993	1	AY236063	Ehrlichia	784	33	82.5	10707	1	AE009321	AE009321 Agrobacte
C 712	33	82.5	993	1	AY236064	Ehrlichia	785	33	82.5	10782	1	AE001639	AE001639 Chlamydia
C 713	33	82.5	993	1	AY236065	Ehrlichia	786	33	82.5	12023	1	AE000876	AE000876 Aquifex a
C 714	33	82.5	993	1	AY236066	Ehrlichia	787	33	82.5	12660	6	AX921008	AX921008 Sequence
C 715	33	82.5	993	1	AY236067	Ehrlichia	788	33	82.5	12702	6	CQ715293	CQ715293 Sequence
C 716	33	82.5	993	1	AY236068	Ehrlichia	789	33	82.5	12719	1	AE010977	AE010977 Methanosa
C 717	33	82.5	993	1	AY236069	Ehrlichia	790	33	82.5	12924	1	AE002059	AE002059 Deinococc
C 718	33	82.5	1017	1	AY455275	Ehrlichia	791	33	82.5	12924	1	AE0081829	AE0081829 Rhiniceph
C 719	33	82.5	1032	11	AY659138	Synthetic	792	33	82.5	14710	2	AE005927	AE005927 Caulobact
C 720	33	82.5	1032	11	AY967083	Synthetic	793	33	82.5	14874	1	AE005927	AE005927 Mus muscu
C 721	33	82.5	1074	6	BD163996	Novel pol	794	33	82.5	15133	9	MMU491857	AE005927 Mus muscu
C 722	33	82.5	1074	6	AX121879	Sequence	795	33	82.5	16343	2	AMFGENOM	L05178 Apis mellif
C 723	33	82.5	1100	6	BD261389	Chlamydia	796	33	82.5	20517	1	AE002181	AE002181 Chlamydo
C 724	33	82.5	1140	8	AY441476	Nycticebu	797	33	82.5	30340	8	AL157899	AL157899 Human DNA
C 725	33	82.5	1149	1	AY236058	Ehrlichia	798	33	82.5	35209	1	AF065159	AF065159 Bradyrhiz
C 726	33	82.5	1197	6	AR477816	Sequence	799	33	82.5	35511	14	AC158189	AC158189 Setaagin
C 727	33	82.5	1197	6	AX066903	Sequence	800	33	82.5	37534	1	AJ872272	AJ872272 Thermotog
C 728	33	82.5	1242	1	AY326315	Streptoco	801	33	82.5	41753	4	AC151390	AC151390 Oryctolag
C 729	33	82.5	1306	1	CHPC820	C.ruminanti	802	33	82.5	42822	8	AF549168	AF549168 Homo sapi
C 730	33	82.5	1316	15	AB182316	AB182316 Heterosig	803	33	82.5	43955	14	AC015342	AC015342 Drosophil
C 731	33	82.5	1332	6	AB320891	Sequence	804	33	82.5	44014	14	AC163895	AC163895 Glomus in
C 732	33	82.5	1431	8	H8U38613	Human marin	805	33	82.5	51811	14	AC100162	AC100162 Mus muscu
C 733	33	82.5	1433	9	BC079060	Rattus no	806	33	82.5	53608	14	AC094555	AC094555 Continuation (4 of
													Continuation (4 of

807	33	82.5	55337	14	AC130377	AC130377 Homo sapi	880	33	82.5	110000	1	BA000008	06	Continuation (7 of
808	33	82.5	61199	14	AC100497	AC100497 Mus muscu	c 881	33	82.5	110000	1	BA000016	18	Continuation (19 o
809	33	82.5	61365	9	AL772192	AL772192 Mouse DNA	c 882	33	82.5	110000	1	BA000028	09	Continuation (10 o
810	33	82.5	62042	14	AP007146	AP007146 Oryza sat	c 883	33	82.5	110000	1	BA000030	58	Continuation (59 o
811	33	82.5	64168	14	AC079003	AC079003 Homo sapi	c 884	33	82.5	110000	1	BA000032	10	BA000032 Vibrio pa
812	33	82.5	66842	14	AC099809	AC099809 Homo sapi	c 885	33	82.5	110000	1	BA000036	17	Continuation (18 o
813	33	82.5	66842	14	AC099809	AC099809 Homo sapi	c 886	33	82.5	110000	1	BA000038	16	Continuation (17 o
814	33	82.5	66984	8	AL606753	AL606753 Human DNA	c 887	33	82.5	110000	1	BA000040	56	Continuation (57 o
815	33	82.5	67152	14	AC099856	AC099856 Mus muscu	c 888	33	82.5	110000	1	BA000045	38	Continuation (39 o
816	33	82.5	67170	14	AC087719	AC087719 Homo sapi	c 889	33	82.5	110000	1	EX571965	29	Continuation (30 o
817	33	82.5	68684	14	AC131271	AC131271 Homo sapi	c 890	33	82.5	110000	1	EX571965	35	Continuation (36 o
818	33	82.5	68790	8	HSJ287H17	AL121970 Human DNA	c 891	33	82.5	110000	1	EX571966	26	Continuation (27 o
819	33	82.5	72147	14	AC021693	AC021693 Homo sapi	c 892	33	82.5	110000	1	EX936398	36	Continuation (37 o
820	33	82.5	72412	14	AC101594	AC101594 Mus muscu	c 893	33	82.5	110000	1	EX950851	22	Continuation (23 o
821	33	82.5	72412	14	AC101594	AC101594 Mus muscu	c 894	33	82.5	110000	1	EX950851	43	Continuation (44 o
822	33	82.5	73408	14	AC130075_3	Continuation (4 of	c 895	33	82.5	110000	1	CP000001	05	Continuation (32 o
823	33	82.5	74049	14	AC135331	AC135331 Homo sapi	c 896	33	82.5	110000	1	CP000010	26	Continuation (6 of
824	33	82.5	74361	9	EX571759	EX571759 Mouse DNA	c 897	33	82.5	110000	1	CP000010	26	Continuation (27 o
825	33	82.5	75001	9	EX678770	EX678770 Mouse DNA	c 898	33	82.5	110000	1	CP000011	01	Continuation (2 of
826	33	82.5	75110	14	AC069115	AC069115 Homo sapi	c 899	33	82.5	110000	1	CP000025	14	Continuation (15 o
827	33	82.5	75689	14	AC018536	AC018536 Homo sapi	c 900	33	82.5	110000	1	CP000031	24	Continuation (25 o
828	33	82.5	77015	14	AC023828	AC023828 Homo sapi	c 901	33	82.5	110000	1	CP000057	12	Continuation (13 o
829	33	82.5	78600	15	SCDNCH2	X79489 S.Cerevisia	c 902	33	82.5	110000	1	CP000058	45	Continuation (46 o
830	33	82.5	81004	14	AC140041	AC140041 Mus muscu	c 903	33	82.5	110000	1	CP000058	58	Continuation (59 o
831	33	82.5	82262	8	AB104726	AB104726 Homo sapi	c 904	33	82.5	110000	1	CP000075	27	Continuation (28 o
832	33	82.5	84011	14	AC018038	AC018038 Drosophil	c 905	33	82.5	110000	1	CP000075	46	Continuation (47 o
833	33	82.5	84108	14	AC133385	AC133385 Rattus no	c 906	33	82.5	110000	1	CP000075	59	Continuation (60 o
834	33	82.5	85159	14	AC166414	AC166414 Bos tauru	c 907	33	82.5	110000	1	CP000076	12	Continuation (11 o
835	33	82.5	86786	14	AC092077	AC092077 Oryza sat	c 908	33	82.5	110000	1	CP000084	10	Continuation (13 o
836	33	82.5	88109	14	AP007661	AP007661 Lotus cor	c 909	33	82.5	110000	1	CP000090	12	Continuation (13 o
837	33	82.5	89211	14	AC133301	AC133301 Rattus no	c 910	33	82.5	110000	6	AR300198	1	Continuation (2 of
838	33	82.5	89703	8	AL356134	AL356134 Human DNA	c 911	33	82.5	110000	6	AR301075	06	Continuation (7 of
839	33	82.5	91110	9	AL928666	AL928666 Mouse DNA	c 912	33	82.5	110000	6	AR607478	06	Continuation (7 of
840	33	82.5	92340	15	AC149267	AC149267 Solanum d	c 913	33	82.5	110000	14	AC092450	5	Continuation (6 of
841	33	82.5	94886	8	HS198G23	AL022150 Human DNA	c 914	33	82.5	110000	14	AC094555	2	Continuation (3 of
842	33	82.5	96554	8	AC004783	AC004783 Homo sapi	c 915	33	82.5	110000	14	AC094997	0	Continuation (3 of
843	33	82.5	96835	14	BA663497_3	Continuation (4 of	c 916	33	82.5	110000	14	AC095248	3	Continuation (4 of
844	33	82.5	98461	15	ATT6H20	AL096859 Arabidops	c 917	33	82.5	110000	14	AC095424	0	AC095424 Rattus no
845	33	82.5	99997	14	AC149997	AC149997 Strongylo	c 918	33	82.5	110000	14	AC095759	0	AC095759 Rattus no
846	33	82.5	100258	8	AC002462	AC002462 Human BAC	c 919	33	82.5	110000	14	AC096839	4	Continuation (5 of
847	33	82.5	100564	15	AC149266	AC149266 Solanum d	c 920	33	82.5	110000	14	AC109746	3	Continuation (4 of
848	33	82.5	101243	8	AC117458	AC117458 Homo sapi	c 921	33	82.5	110000	14	AC111624	2	Continuation (3 of
849	33	82.5	102352	1	BA000004_41	Continuation (42 o	c 922	33	82.5	110000	14	AC112426	0	AC112426 Rattus no
850	33	82.5	103022	15	AP003909	AP003909 Oryza sat	c 923	33	82.5	110000	14	AC114091	0	AC114091 Rattus no
851	33	82.5	103300	14	AL592542	AL592542 Danio rer	c 924	33	82.5	110000	14	AC114829	1	Continuation (2 of
852	33	82.5	103488	9	AL772399	AL772399 Mouse DNA	c 925	33	82.5	110000	14	AC115558	3	Continuation (4 of
853	33	82.5	104081	8	HSJ263J7	AL049545 Human DNA	c 926	33	82.5	110000	14	AC118415	0	AC118415 Rattus no
854	33	82.5	106278	14	RM520N23	AL732853 Rattus no	c 927	33	82.5	110000	14	AC118415	1	Continuation (2 of
855	33	82.5	106938	8	AC132032	AC132032 Homo sapi	c 928	33	82.5	110000	14	AC118415	2	Continuation (3 of
856	33	82.5	109790	14	AC129739	AC129739 Rattus no	c 929	33	82.5	110000	14	AC130075	2	Continuation (3 of
857	33	82.5	110000	1	AE002098_02	Continuation (3 of	c 930	33	82.5	110000	14	AC130075	2	Continuation (3 of
858	33	82.5	110000	1	CP000091_09	Continuation (10 o	c 931	33	82.5	110000	14	EX663497_1	1	Continuation (2 of
859	33	82.5	110000	1	CP000107_11	Continuation (12 o	c 932	33	82.5	110000	14	CT005266	00	CT005266 Leishmani
860	33	82.5	110000	1	CR626927_27	Continuation (28 o	c 933	33	82.5	110000	14	LMFLCHR32_20	20	Continuation (21 o
861	33	82.5	110000	1	CR626927_28	Continuation (29 o	c 934	33	82.5	110000	15	AP008214_002	002	Continuation (3 of
862	33	82.5	110000	1	CR767821_13	Continuation (14 o	c 935	33	82.5	110000	15	AP008214_221	221	Continuation (222
863	33	82.5	110000	1	CR925677_13	Continuation (14 o	c 936	33	82.5	110000	15	AP008215_146	146	Continuation (147
864	33	82.5	110000	1	CR925678_13	Continuation (14 o	c 937	33	82.5	110000	15	AP008215_176	176	Continuation (177
865	33	82.5	110000	1	AB000951_05	Continuation (6 of	c 938	33	82.5	110000	15	AP008215_196	196	Continuation (197
866	33	82.5	110000	1	AE016822_07	Continuation (8 of	c 939	33	82.5	110000	15	AP008215_197	197	Continuation (198
867	33	82.5	110000	1	AE016853_47	Continuation (48 o	c 940	33	82.5	110000	15	AP008215_213	213	Continuation (214
868	33	82.5	110000	1	AE016853_62	Continuation (63 o	c 941	33	82.5	110000	15	CR380957_10	10	Continuation (11 o
869	33	82.5	110000	1	AE017225_31	Continuation (32 o	c 942	33	82.5	110000	15	CR382135_05	05	Continuation (6 of
870	33	82.5	110000	1	AE017334_31	Continuation (32 o	c 943	33	82.5	110000	15	CR382139_02	02	Continuation (3 of
871	33	82.5	110000	1	AE017355_31	Continuation (32 o	c 944	33	82.5	110000	15	AE017348_01	01	Continuation (2 of
872	33	82.5	110000	1	AE017355_32	Continuation (33 o	c 945	33	82.5	110000	15	AP008209_193	193	Continuation (194
873	33	82.5	110000	1	AJ749949_03	Continuation (4 of	c 946	33	82.5	110000	15	AP008209_236	236	Continuation (237
874	33	82.5	110000	1	AJ749949_05	Continuation (6 of	c 947	33	82.5	110000	15	AP008210_080	080	Continuation (81 o
875	33	82.5	110000	1	AP006716_20	Continuation (21 o	c 948	33	82.5	110000	15	AP008210_217	217	Continuation (218
876	33	82.5	110000	1	AP008226_05	Continuation (6 of	c 949	33	82.5	110042	8	AL662904	4	AL662904 Human DGA
877	33	82.5	110000	1	BA000001_07	Continuation (8 of	c 950	33	82.5	111089	15	AC146758	15	AC146758 Medicago
878	33	82.5	110000	1	BA000004_10	Continuation (11 o	c 951	33	82.5	117328	6	CQ870156	6	CQ870156 Sequence
879	33	82.5	110000	1	BA000004_11	Continuation (12 o	c 952	33	82.5	117573	14	AP007654	14	AP007654 Lotus cor


```

c 953 33 82.5 118635 15 AP005862
c 954 33 82.5 119044 14 AC018579
c 955 33 82.5 119743 14 AC134515
c 956 33 82.5 120199 8 AP003498
c 957 33 82.5 120910 8 AC105269
c 958 33 82.5 121543 14 AL772350
c 959 33 82.5 121613 14 AC149977
c 960 33 82.5 123471 5 BX640408
c 961 33 82.5 123897 14 AC144684
c 962 33 82.5 124240 15 OSJN00042
c 963 33 82.5 125354 14 AC155894
c 964 33 82.5 125478 14 AC154082
c 965 33 82.5 127347 14 AP000682
c 966 33 82.5 127482 5 CR407560
c 967 33 82.5 127549 14 AC155384
c 968 33 82.5 127724 14 AC160096
c 969 33 82.5 128155 8 HSA262A13
c 970 33 82.5 128361 8 AC004877
c 971 33 82.5 128458 15 AC153125
c 972 33 82.5 129361 8 AL672292
c 973 33 82.5 129610 8 AP006212
c 974 33 82.5 130127 4 AC153760
c 975 33 82.5 130263 8 HS681N20
c 976 33 82.5 131465 8 AL731573
c 977 33 82.5 131570 8 AP002495
c 978 33 82.5 131117 5 AL732636
c 979 33 82.5 134222 8 AB045363
c 980 33 82.5 134757 8 AC007870
c 981 33 82.5 135011 14 AC020183
c 982 33 82.5 135079 8 HSDJ70A9
c 983 33 82.5 136502 8 HS705F19
c 984 33 82.5 136560 9 AC102475
c 985 33 82.5 136754 14 GR394528
c 986 33 82.5 136893 14 AC136643
c 987 33 82.5 136893 14 AC136643
c 988 33 82.5 137525 15 AP005397
c 989 33 82.5 137628 9 AL672153
c 990 33 82.5 137905 14 AC015605
c 991 33 82.5 138126 14 AC108760
c 992 33 82.5 138313 8 AC092111
c 993 33 82.5 139190 9 AC139329
c 994 33 82.5 139224 8 AC093857
c 995 33 82.5 140393 14 AC141097
c 996 33 82.5 140584 5 BX510911
c 997 33 82.5 141674 14 AC080046
c 998 33 82.5 141870 14 AC092986
c 999 33 82.5 141915 15 OSJN00024
1000 33 82.5 142045 14 AC141650

ALIGNMENTS

RESULT 1
CQ687716 230 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 32642 from Patent WO02070737.
DEFINITION CQ687716
ACCESSION CQ687716
VERSION CQ687716.1 GI:42218962
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Liew, C.C., Marshall, W.E. and Zhang, H.
Compositions and methods relating to osteoarthritis
Patent: WO 02070737-A 32642 12-SEP-2002;
Chondrogene Inc. (CA)
LOCATION/Qualifiers
1..290
/organism="Homo sapiens"
/mol_type="unassigned DNA"

AP005862 Oryza sat
AC018579 Homo sapi
AC134515 Rattus no
AP003498 Homo sapi
AC105269 Homo sapi
AL772350 Mus muscu
BX640408 Zebrafish
AC144684 Rattus no
AL605597 Oryza sat
AC155894 Medicago
AC154082 Dromaius
AP000682 Homo sapi
CR407560 Zebrafish
AC155384 Zee mayas
AC160096 Medicago
AL096843 Human DNA
AC004877 Homo sapi
AC153125 Medicago
AL672292 Human DNA
AP006212 Homo sapi
AC153760 Loxodonta
AL031670 Human DNA
AL731573 Human DNA
AP002495 Homo sapi
AL732636 Tetraodon
AB045363 Homo sapi
AC007870 Genomic s
AC020183 Drosoephil
AL121788 Human DNA
AL035415 Human DNA
AC102475 Mus muscu
CR394528 Danio rer
AC136643 Rattus no
AC136643 Rattus no
AP005397 Oryza sat
AL672153 Mouse DNA
AC015605 Mus muscu
AC108760 Oryza sat
AC092111 Homo sapi
AC139329 Mus muscu
AC093857 Homo sapi
AC141097 Rattus no
BX510911 Zebrafish
AC080046 Homo sapi
AC092986 Homo sapi
AL605592 Oryza sat
AC141650 Rattus no

/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 4.74 Length: 230
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-6 (1-9) x CQ687716 (1-290)
Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
|||||
Db 71 GCCCTGATAGGGGCTATTTCCTCTG 97

RESULT 2
CQ920916 475 bp DNA linear PAT 23-NOV-2004
LOCUS Sequence 2116 from Patent WO2004097052.
DEFINITION CQ920916
ACCESSION CQ920916
VERSION CQ920916.1 GI:56210857
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Burczynski, M.E., Twine, N.C., Slonim, D.K., Trepicchio, W.L.,
Strahs, A., Immerman, P. and Dornier, A.J.
Methods for prognosis and treatment of solid tumors
Patent: WO 2004097052-A 2116 11-NOV-2004;
Wyeth (US); Burczynski, Michael E. (US)
LOCATION/Qualifiers
1..475
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 8.11 Length: 475
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-6 (1-9) x CQ920916 (1-475)
Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
|||||
Db 369 GCCCTGATAGGGGCTATTTCCTCTG 395

RESULT 3
BD249733 901 bp DNA linear PAT 17-JUL-2003
LOCUS Polypeptide.
DEFINITION BD249733
ACCESSION BD249733
VERSION BD249733.1 GI:33059503
KEYWORDS JP 2002530060-A/3.
SOURCE Canis sp.
ORGANISM Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
1 (bases 1 to 901)
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 3 17-SEP-2002;
OXFORD BIOMEDICA LTD
```



```
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 16.8 Length: 927
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX829164 (1-927)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 748 GCCCTGATAGCGCTATTTCCTCTG 774

RESULT 7
AX821533
LOCUS AX821533 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO03068812.
ACCESSION AX821533
VERSION AX821533.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"

ORIGIN
Alignment Scores:
Pred. No.: 23.5 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX467373 (1-1260)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1087 GCCCTGATAGTGCCATTTCCTACTG 1113

RESULT 8
AX821533
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Carroll, M.M., Kingsman, S.M. and Redchenko, I.M.
TITLE MHC class I peptide epitopes from the human 5t4 tumor-associated
antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;

/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN
Alignment Scores:
Pred. No.: 23.5 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX821533 (1-1260)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1087 GCCCTGATAGTGCCATTTCCTACTG 1113

RESULT 9
AX821548
LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Carroll, M.O., Harrop, R.O. and Kingsman, S.O.
TITLE MHC class II peptide epitope of 5t4 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN
Alignment Scores:
Pred. No.: 23.5 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX821548 (1-1260)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1087 GCCCTGATAGTGCCATTTCCTACTG 1113

RESULT 10
BD249731
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 (bases 1 to 1263)
```

AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL,KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
PC C07K13/00,
PC C12N15/00
CC Polypeptide
FH Key Location/Qualifiers
FT source 1..1263
FT Location/Qualifiers
source 1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 23.5 Length: 1263
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-6 (1-9) x BD249731 (1-1263)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 1090 GCCCTGATAGGCGCTATTTCTCCTCG 1116
RESULT 11
AX025011
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO0029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source 1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 23.5 Length: 1263
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-6 (1-9) x AX025011 (1-1263)

AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL,KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
PC C07K13/00,
PC C12N15/00
CC Polypeptide
FH Key Location/Qualifiers
FT source 1..1263
FT Location/Qualifiers
source 1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 23.5 Length: 1263
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-6 (1-9) x BD249731 (1-1263)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 1090 GCCCTGATAGGCGCTATTTCTCCTCG 1116
RESULT 11
AX025011
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO0029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source 1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 23.5 Length: 1263
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-6 (1-9) x AX025011 (1-1263)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 1090 GCCCTGATAGGCGCTATTTCTCCTCG 1116
RESULT 12
AX149553
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
Ellard,F.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="574"
ORIGIN
Alignment Scores: 23.5 Length: 1263
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-6 (1-9) x AX149553 (1-1263)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 1090 GCCCTGATAGGCGCTATTTCTCCTCG 1116
RESULT 13
AX316086
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:117899278
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 23.5 Length: 1263
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6
US-10-774-176-6 (1-9) x AX316086 (1-1263)

```
DB:                                     0          Gaps:
US-10-774-176-6 (1-9) x AX316086 (1-1263)
Qy 1 AlaLeuIlleGlyAlaIlePheLeuLeu 9
Db 1090 GCCCTGATAGGCGCTATTTCCTCTG 1116

RESULT 14
AX467371 AX467371 1263 bp DNA linear PAT 16-JUL-2002
LOCUS Sequence 1 from Patent WO0238612.
DEFINITION AX467371
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS Canis sp.
SOURCE Canis sp.
ORGANISM Canis sp.
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1263
Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN
Alignment Scores:
Pred. No.: 23.5 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-774-176-6 (1-9) x AX467371 (1-1263)
Qy 1 AlaLeuIlleGlyAlaIlePheLeuLeu 9
Db 1090 GCCCTGATAGGCGCCATCTTCTACTG 1116

RESULT 15
BD249732 BD249732 1281 bp DNA linear PAT 17-JUL-2003
LOCUS Polypeptide.
DEFINITION BD249732
ACCESSION BD249732
VERSION BD249732.1 GI:33059502
KEYWORDS JP 2002530060-A/2.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1281)
Carroll,M.W. and Myers,K.A.
Polypeptide
Patent: JP 2002530060-A 2 17-SEP-2002;
OXFORD BIOMEDICA LTD
OS Mus musculus (mouse)
PN JP 2002530060-A/2
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09, A61K39/00, A61K48/00, C07K7/06, C07K14/065,
PC C07K19/00,
PC C12N15/00

CC Polypeptide Location/Qualifiers
FH Key 1..1281
FT source /organism="Mus musculus (mouse)".

FEATURES
source
1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 23.9 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-774-176-6 (1-9) x BD249732 (1-1281)
Qy 1 AlaLeuIlleGlyAlaIlePheLeuLeu 9
Db 1108 GCTCTGATAGGCGCTATTTCCTCTC 1134

RESULT 16
AX025012 AX025012 1281 bp DNA linear PAT 15-SEP-2000
LOCUS Sequence 2 from Patent WO0029428.
DEFINITION AX025012
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
Carroll,M.W. and Myers,K.A.
Polypeptide
Patent: WO 0029428-A 2 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
Location/Qualifiers
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

FEATURES
source
1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 23.9 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-774-176-6 (1-9) x AX025012 (1-1281)
Qy 1 AlaLeuIlleGlyAlaIlePheLeuLeu 9
Db 1108 GCTCTGATAGGCGCTATTTCCTCTC 1134

RESULT 17
AX316087 AX316087 1281 bp DNA linear PAT 14-DEC-2001
LOCUS Sequence 2 from Patent EP160323.
DEFINITION AX316087
ACCESSION AX316087
VERSION AX316087.1 GI:17899279
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE 5T4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 2 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 23.9 Length: 1281
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX316087 (1-1281)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1108 GCTCTGATAGGCGCTATTTCCTCCTC 1134

RESULT 18
CQ731678
LOCUS CQ731678 2053 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 17612 from Patent WO02069579.
ACCESSION CQ731678
VERSION CQ731678.1 GI:42308932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 17612 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
source
1..2053
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 40 Length: 2053
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x CQ731678 (1-2053)

Qy 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1176 GCCCTGATAGGCGCTATTTCCTCCTG 1202

RESULT 19
HS5T4OA
LOCUS HS5T4OA 2053 bp RNA linear PRI 18-APR-2005
DEFINITION Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.

229083
229083.1 GI:435654
5T4 gene: 5T4 oncofoetal antigen.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2053)
AUTHORS Myers, K.A., Rahi-Saund, V., Davison, M.D., Young, J.A., Cheater, A.J.
and Stern, P.L.
TITLE Isolation of a cDNA encoding 5T4 oncofoetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
JOURNAL J. Biol. Chem. 269 (12), 9319-9324 (1994)
PUBMED 8132670
REFERENCE
2 (bases 1 to 2053)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
FEATURES
Location/Qualifiers
1..2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label=N-flank
85..1347
/codon_start=1
/evidence=experimental
/product="5T4 oncofoetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGSGPAGDGRRLRLARLALVLGWSSSPSTSSASFSS
APFLASAVQAQPLPDQCPALCESEARTKVCNRLTEVPTDLDPAAYVRLFLTGQ
LAVLPAGAPARPLAELALNLGSLRDLVRAGAPEHLPLSLRQLDLSHNLPLDLSPP
APSGNASVSAPSPVLVLIHLNIVPPEDRQNEPSFGMVVAALLAGLQGLRLELA
SNHFLYLPDVLQALPSLRHLDSLNSLSVLTYSVFNLTLSLHLEDNALKVHLNG
TLAEIQGLPHIRVFLDNNPWCDCMADVTMLKETEVVGQKDRLTCAYPEKMRNRL
LELNGADLDCDPILPESLQTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMHIRDAC
RDHMEGYHYRYEINADPRLTNLSSNDV"
130..171
/products="Leucine rich repeat region"
373..966
/label=LRRS
966..1119
/product="LRR C-terminal flank"
/label=C-flank
1153..1215
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

sig_peptide
misc_RNA
373..966
/products="Leucine rich repeat region"
/label=LRRS
966..1119
/product="LRR C-terminal flank"
/label=C-flank
1153..1215
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

ORIGIN
Alignment Scores:
Pred. No.: 40 Length: 2053
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

```
US-10-774-176-6 (1-9) x HS5T4OA (1-2053)

Qy 1 AlaleuileGlyAlaIlePheLeu 9
Db 1174 GCCCTGATAGCGCTATTTCCTCTG 1200

RESULT 20
AF063939 Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA, ROD 01-JAN-2000
LOCUS complete cds.
DEFINITION Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA, ROD 01-JAN-2000
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 2333)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
TITLE Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2333)
AUTHORS Ninkina, N.N. and Buchman, V.L.
TITLE Structure and expression of the rat 5T4 gene
JOURNAL Direct Submission
SUBMITTED (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK
FEATURES
source 1..2333
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="newborn"
1..2333
/gene="5T4"
1..363
/gene="5T4"
364..1644
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal antigen homolog"
/protein_id="AAF21770.1"
/db_xref="GI:6650212"
/translaton="MPGAGSRGPGAGDGRGLRLALVLLGWVSASAPSSLPSSSTS
PAATLASAOPPAERCPAAECSEAAATVKVNRNLLEVPADLPVYRNLFLTGNO
MTVLPAAGAFAPQPLADLAVLNLGNHLKEVGAGAFELPGLRLDLSHNPILTNSAP
TFAGNSVSTPSPLELILNHIIVPPDQORNGFEGVAFEGWAAALRGLALRGL
HHLASNLHFLYLPRLDLDQLPSLKHLDLRNLSVLTYSFRNLTHLSLHLSDNAL
KVLHNSLAEVQGLARVRFVLDNNPWCDCYMDVMVSLKTEVVPDKARLTCAFPPEK
MRNRLDLTSLSDCCDCTLPSQLQTSYVFLGIVLALIGALFLVLLVLRNKGKXKWH
NIRDACRDHMEGYHYREINADPSLTNLSNSDV"
1645..2333
/gene="5T4"
2315..2320
/gene="5T4"

3'UTR
polyA_signal

ORIGIN
Alignment Scores: 45.9 Length: 2333
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 9

US-10-774-176-6 (1-9) x AF063939 (1-2333)

Qy 1 AlaleuileGlyAlaIlePheLeu 9
Db 1471 GCTCTGATAGCGCTATTTCCTCTC 1497

US-10-774-176-6 (1-9) x HS5T4OA (1-2053)

Qy 1 AlaleuileGlyAlaIlePheLeu 9
Db 1174 GCCCTGATAGCGCTATTTCCTCTG 1200

RESULT 20
AF063939 Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA, ROD 01-JAN-2000
LOCUS complete cds.
DEFINITION Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA, ROD 01-JAN-2000
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 2333)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
TITLE Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2333)
AUTHORS Ninkina, N.N. and Buchman, V.L.
TITLE Structure and expression of the rat 5T4 gene
JOURNAL Direct Submission
SUBMITTED (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK
FEATURES
source 1..2333
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="newborn"
1..2333
/gene="5T4"
1..363
/gene="5T4"
364..1644
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal antigen homolog"
/protein_id="AAF21770.1"
/db_xref="GI:6650212"
/translaton="MPGAGSRGPGAGDGRGLRLALVLLGWVSASAPSSLPSSSTS
PAATLASAOPPAERCPAAECSEAAATVKVNRNLLEVPADLPVYRNLFLTGNO
MTVLPAAGAFAPQPLADLAVLNLGNHLKEVGAGAFELPGLRLDLSHNPILTNSAP
TFAGNSVSTPSPLELILNHIIVPPDQORNGFEGVAFEGWAAALRGLALRGL
HHLASNLHFLYLPRLDLDQLPSLKHLDLRNLSVLTYSFRNLTHLSLHLSDNAL
KVLHNSLAEVQGLARVRFVLDNNPWCDCYMDVMVSLKTEVVPDKARLTCAFPPEK
MRNRLDLTSLSDCCDCTLPSQLQTSYVFLGIVLALIGALFLVLLVLRNKGKXKWH
NIRDACRDHMEGYHYREINADPSLTNLSNSDV"
1645..2333
/gene="5T4"
2315..2320
/gene="5T4"

3'UTR
polyA_signal

ORIGIN
Alignment Scores: 45.9 Length: 2333
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 9

US-10-774-176-6 (1-9) x AF063939 (1-2333)

Qy 1 AlaleuileGlyAlaIlePheLeu 9
Db 1471 GCTCTGATAGCGCTATTTCCTCTC 1497
```

```
RESULT 21
BD127282
LOCUS 2359 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127282
VERSION BD127282.1 GI:23222227
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otauki, T. and
Koga, H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/2713
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUO OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ 10,
C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
(424)..(1572).
1..2359
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 46.5 Length: 2359
Pred. No.: 40.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 6

US-10-774-176-6 (1-9) x BD127282 (1-2359)

Qy 1 AlaleuileGlyAlaIlePheLeu 9
Db 1513 GCCCTGATAGCGCTATTTCCTCTG 1539

RESULT 22
CQ782724
LOCUS 2359 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2864 from Patent EP1396543.
ACCESSION CQ782724
VERSION CQ782724.1 GI:45502667
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otauki, T. and
Koga, H.
```

```

TITLE
JOURNAL
Primers for synthesizing full length cDNA clones and their use
Patent: EP 1396543-A 2864 10-MAR-2004;
Research Association for Biotechnology (JP)
FEATURES
source
Location/Qualifiers
1..2359
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"
/translation="MPGCGSRGPAAGDGLRLRLALVLGWSSSSPTSSASSPSSS
APFLASVASQPLPDQPCALCESEARTVKYNRLTEPTDLPAYVRLPLTGNQ
LAVPAGFAARPLPLAELNLGSLRDEVRAGAFELHPLSLRQLDLGHNPLADLSPF
APSGNSAASPPSPLEVLILNIHVPDEQRNRPSEGMVAALAGRALQGLRLLELA
SHFLYLPRDYLQAQLDLSNLSVLTYSFRNLTHLESLELDNALKVLHNG
TLAEQLGLPHRVFLDNNPWCDCMDMVTWLKETEYVQGRDLTCAYPEKMRNRL
LELNSADLDCDPLPPSLQTSVFLGIVLALIGALFLVLVLYLRNGIKK"
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424..1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"
/translation="MPGCGSRGPAAGDGLRLRLALVLGWSSSSPTSSASSPSSS
APFLASVASQPLPDQPCALCESEARTVKYNRLTEPTDLPAYVRLPLTGNQ
LAVPAGFAARPLPLAELNLGSLRDEVRAGAFELHPLSLRQLDLGHNPLADLSPF
APSGNSAASPPSPLEVLILNIHVPDEQRNRPSEGMVAALAGRALQGLRLLELA
SHFLYLPRDYLQAQLDLSNLSVLTYSFRNLTHLESLELDNALKVLHNG
TLAEQLGLPHRVFLDNNPWCDCMDMVTWLKETEYVQGRDLTCAYPEKMRNRL
LELNSADLDCDPLPPSLQTSVFLGIVLALIGALFLVLVLYLRNGIKK"
ORIGIN
Alignment Scores:
Pred. No.: 46.5 Length: 2359
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-6 (1-9) x CQ782724 (1-2359)
Qy 1 AlaLeuIlleGlyAlaIlePheLeuLeu 9
Db 1513 GCCCTGATAGGCGCTATTTCCTCCTG 1539
RESULT 23
AK074786 2359 bp mRNA linear PRI 03-SEP-2002
LOCUS
DEFINITION
Homo sapiens cDNA FLJ90305 fis, clone NT2RP2000694, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION
AK074786
VERSION
AK074786.1 GI:22760460
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masuo,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEBO human cDNA sequencing project
Unpublished
2 (bases 1 to 2359)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1..2359
FEATURES
source

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"
ORIGIN
Alignment Scores:
Pred. No.: 46.5 Length: 2359
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-10-774-176-6 (1-9) x AK074786 (1-2359)
Qy 1 AlaLeuIlleGlyAlaIlePheLeuLeu 9
Db 1513 GCCCTGATAGGCGCTATTTCCTCCTG 1539
RESULT 24
BD127283 2361 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD127283
VERSION
BD127283.1 GI:23222228
KEYWORDS
JP 2002017375-A/2714.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2361)
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 2714 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2714
PD 22-JAN-2002 JP 20020253172
PF 07-JUL-2000 JP 20020253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
PC
C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers
(426)..(1685).
FT CDS
FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 46.5 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0

```


Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x BD127283 (1-2361)

Qy 1 AlaleulleGlyAlaIlePheLeuLeu 9
|||||
Db 1515 GCCGTAGAGGCGCTATTTCCTCTG 1541

RESULT 25
CQ782726 2361 bp DNA linear PAT 17-MAR-2004
LOCUS
DEFINITION Sequence 2866 from Patent EP1396543.
ACCESSION CQ782726
VERSION CQ782726.1 GI:45502669
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2866 10-MAR-2004;
Research Association for Biotechnology (JP)

FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
426..1688
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85961.1"
/db_xref="GI:45502670"

CDS
46.5 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX961916 (1-2361)

Qy 1 AlaleulleGlyAlaIlePheLeuLeu 9
|||||
Db 1515 GCCGTAGAGGCGCTATTTCCTCTG 1541

RESULT 27
AK074790 2361 bp mRNA linear PRI 09-JUL-2005
LOCUS
DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AK074790
VERSION AK074790.1 GI:22760466
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS Otsuki, T., Ota, T., Nishikawa, T., Hayashi, K., Suzuki, Y.,
Yamamoto, J., Wakamatsu, A., Kimura, K., Sakamoto, K., Hatanoto, N.,
Kawai, Y., Ishii, S., Saito, K., Kojima, S., Sugiyama, T., Ono, T.,
Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A.,
Okumura, K., Nagai, K., Sugano, S. and Isogai, T.
TITLE Signal Sequence and Keyword Trap in silico for Selection of
Full-Length Human cDNAs Encoding Secretion or Membrane Proteins
from Oligo-Capped cDNA Libraries
JOURNAL DNA Res. 12, 117-126 (2005)
REFERENCE
AUTHORS Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahari, K., Maguho, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2361)

Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x BD127283 (1-2361)

Qy 1 AlaleulleGlyAlaIlePheLeuLeu 9
|||||
Db 1515 GCCGTAGAGGCGCTATTTCCTCTG 1541

RESULT 25
CQ782726 2361 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 127 from Patent W003104277.
ACCESSION AX961916
VERSION AX961916.1 GI:40881326
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2866 10-MAR-2004;
Research Association for Biotechnology (JP)

FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
426..1688
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85961.1"
/db_xref="GI:45502670"

CDS
46.5 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x CQ782726 (1-2361)

Qy 1 AlaleulleGlyAlaIlePheLeuLeu 9
|||||
Db 1515 GCCGTAGAGGCGCTATTTCCTCTG 1541

RESULT 26
AX961916 2361 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 127 from Patent W003104277.
ACCESSION AX961916
VERSION AX961916.1 GI:40881326
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1..2361
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="Cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"
ORIGIN
Alignment Scores:
Pred. No.: 46.5 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatve: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-10-774-176-6 (1-9) x AK074790 (1-2361)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 1515 GCCCTGATAGCGCTATTTCTCTCTG 1541
RESULT 28
BC087011
LOCUS Rattus norvegicus trophoblast glycoprotein, mRNA (cDNA clone
MGC:93332 IMAGE:7193411), complete cds.
ACCESSION BC087011.1 GI:56268819
VERSION MGC.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2361)
AUTHORS Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S.S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length

JOURNAL PUBMED
REFERENCE 2 (bases 1 to 2361)
AUTHORS Director MGC Project.
TITLE Submitted
JOURNAL Submitted (02-DEC-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
CDNA Sequencing by: The I.M.A.G.E. Consortium (ILLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-ahgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov
Series: IRAC Plate: 186 Row: 0 Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13929143.
FEATURES
source
1..2361
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93332 IMAGE:7193411"
/tissue_type="Heart, rat (Brown Norway)"
/clone_lib="NIH MGC_234"
/lab_host="DH10B"
/note="Vector: pExpress1"
1..2361
/gene="Tpbpg"
/note="synonym: 5T4"
/db_xref="GeneID:83684"
/db_xref="RGD:621453"
364..1644
/gene="Tpbpg"
/codon_start=1
/product="Tpbpg protein"
/protein_id="AAH87011.1"
/db_xref="GI:56268820"
/db_xref="GeneID:83684"
/db_xref="RGD:621453"
/translation="MPGAGSRGPSAGDGRRLRLRLALVLLGWVSAGAPSSSLPSSSTS
PAAFLASGAQPPPAERCPAACESEARTVKCNVNLLEVPADLPYVNLPLTGNQ
MTVLPAGAFARQPLADLVNLISGNHLKVEGAGAFELPLRLDLSHNPLNLSAF
TFAGNSVSTPSPLELNLHNIVPPEDQRNGSPGVAFGWAAALRSGLALGL
HLELASLHLYPLRDLQLPSLKHLDLNNLSVLTYSFNLTHLSLHLELNLAL
KVLNLTAAEQGLAHVPLDNLNPMWVCYMDVMYSLKTEVVPDKARLTCAFPK
MENRGLDITSSDLDCDADTLPSQSLQTSYVFLGIVLALIGALFLVLYLNKGIKKWMH
NTRDACRDHMEGVHYREINADPRNLNLSNSDV"
ORIGIN
Alignment Scores:
Pred. No.: 46.5 Length: 2361
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatve: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0
US-10-774-176-6 (1-9) x BC087011 (1-2361)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
|||||

```

Db      1471 GCTCTGATAGGCGCTATTTCCTCTC 1497

RESULT 29
BC037161
LOCUS   Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317
DEFINITION IMAGE:4138906), complete cds.
ACCESSION BC037161
VERSION   BC037161.2 GI:33872201
KEYWORDS  MGC.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo
REFERENCE 1 (bases 1 to 2379)
AUTHORS  Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
          Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
          Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Schenck, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
          Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
          Worsley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
          Villaalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodrigues, S.,
          Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalish, D.E.,
          Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL 12477932
PUBMED 2 (bases 1 to 2379)
REFERENCE Strausberg, R.
AUTHORS   Direct Submission
TITLE     Submitted (03-SEP-2002) National Institutes of Health, Mammalian
JOURNAL  Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT  On Aug 19, 2003 this sequence version replaced gi:22713382.
          Contact: MGC help desk
          Email: cgapsb-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland;
          Web site: http://www.nisc.nih.gov/
          Contact: nisc.mgc@nri.nih.gov
          Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
          Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
          Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
          Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
          Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
          McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
          Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
          Young, A., Zhang, L.-H. and Green, E.D.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
          Series: IRAL Plate: 26 Row: m Column: 15
          This clone was selected for full length sequencing because it
          passed the following selection criteria: matched mRNA gi: 5729717.
          Location/Qualifiers
          1. .2379

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:15317 IMAGE:4138906"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. .2379
/gene="TPBG"
/note="Synonyms: M6P1, 5T4-AG, 5T4"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"
427..1689
/gene="TPBG"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="AAH37161.1"
/db_xref="GI:22713383"
/db_xref="GeneID:7162"
/db_xref="MIM:190920"
/translation="MPGCSGRPAAGDGRRLRLRLALVLLGWSSSPSSASSPSS
APFLASVAQPPPLDQPCALCESEARTVKVNRNLTEVPTDLPAYVNRNLFITGNQ
LAVLPAGAFARPPPLAEALNLSGRSLDEVRAGAFELPSLRQLDLSNPLADLSFP
AFSGNSASVAPSPIVELIHNHIVPDERQNRSPFGMVVAALAGRALQGLRLLELA
SNHFLYLPDVLQALPSLRHLDSNLSVSLTVSVFNLTHLESIHLEADNALKVHLNG
THLEQGLPHRVFLDNNPVWCDCHMADMTWLKETEVOGKDRLTCAYPEKMRNVL
LELSADDDCDPILPSPQTSVTFGLVILALIGALFLVLYLNRGKIKKMMNIRDAC
RDHMEGYHYRYEINADPRLTNLSSNDV"
Alignment Scores:
Pred. No.: 45.9 Length: 2379
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0
US-10-774-176-6 (1-9) x BC037161 (1-2379)
QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
|||||
DB 1516 GCCTGATAGGCGCTATTTCCTCTG 1542
BC058198 2423 bp mRNA linear ROD 21-OCT-2003
Mus musculus trophoblast glycoprotein, mRNA (cDNA clone MGC:68145
IMAGE:5353871), complete cds.
BC058198
MGC.
BC058198.1 GI:34849573
Mus musculus (house mouse)
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2423)
AUTHORS  Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
          Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Schenck, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
          Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,
          Worsley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
          Villaalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodrigues, S.,
          Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalish, D.E.,
          Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL 12477932
PUBMED 2 (bases 1 to 2379)
REFERENCE Strausberg, R.
AUTHORS   Direct Submission
TITLE     Submitted (03-SEP-2002) National Institutes of Health, Mammalian
JOURNAL  Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT  On Aug 19, 2003 this sequence version replaced gi:22713382.
          Contact: MGC help desk
          Email: cgapsb-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland;
          Web site: http://www.nisc.nih.gov/
          Contact: nisc.mgc@nri.nih.gov
          Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
          Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
          Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
          Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
          Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
          McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
          Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
          Young, A., Zhang, L.-H. and Green, E.D.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
          Series: IRAL Plate: 26 Row: m Column: 15
          This clone was selected for full length sequencing because it
          passed the following selection criteria: matched mRNA gi: 5729717.
          Location/Qualifiers
          1. .2379

```

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2423)
 Strausberg,R.
 Direct Submission
 Submitted (15-SEP-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Grant,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
 Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
 Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 123 Row: P Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

Location/Qualifiers
 1..2423
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:68145 IMAGE:5353871"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH108"
 /note="Vector: pCMV-SPORT6"

gene

1..2423
 /gene="Tpbp"
 /note="synonym: 574"
 /db_xref="GeneID:574"
 /db_xref="MGI:1341264"
 402..1682

CDS

/gene="Tpbp"
 /codon_start=1
 /product="Tpbp protein"
 /protein_id="AAH58198.1"
 /db_xref="GI:34849574"
 /db_xref="GeneID:21983"
 /db_xref="MGI:1341264"
 /translation="MPCAGSRGSPAGDRLRLARLALVLLGWVSAPSSVPSSSTS
 PAAFLASGAQPPPAERCPACCEAARTKVCNRNLLLEVPADLPYVNRNLFITGNO
 MTVLPAGAFARQPPADLALNLSGNHLKEVCAGAFELPGLRLDLSHNPLTNLSAF
 APAGSNASVSPLEELILNHIVPPEDQKQSGFEGMVAFGMVAALRSGLALRGL
 TCLEASHNHFELPDLRLAQLPSRLYLDLRNNSLSVLTYSFRNLTHLESILHLEDNAL
 KVLHNSTLAEWQGLAHVKVFLDNNPWCDCYMDADMWALVLTETVVPDKARLTCAPEK
 MNRNGLDLNSDLDCDAVLPSQLSQTSYVFLGIVLALIGAIFLLVLYLNRRKGIKQWH
 NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"

misc_feature
 642..1262
 /gene="Tpbp"
 /note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
 protein [function unknown]"
 /db_xref="CDD:COG4886"
 misc_feature
 1299..1415
 /gene="Tpbp"
 /note="LRRCT; Region: Leucine rich repeat C-terminal
 domain"
 /db_xref="CDD:smart00082"

ORIGIN

Alignment Scores:
 Pred. No.: 47.9 Length: 2423
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-6 (1-9) x BC058198 (1-2423)

QY 1 AlaLeuileglyAlaIlePheLeuLeu 9

Db 1509 GCTCTGATAGGCGCTATTTCTCTCTC 1535

RESULT 31

AX961912
 LOCUS AX961912 2557 bp DNA linear PAT 14-JAN-2004
 DEFINITION Sequence 123 from Patent WO03104277.
 AX961912
 ACCESSION AX961912 GI:40881322

VERSION

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Location/Qualifiers

1..2557

/organism="Mus musculus"

/mol_type="unassigned DNA"

/db_xref="taxon:10090"

556..1836

/note="unnamed protein product"

/codon_start=1

/protein_id="CAF06465.1"

/db_xref="GI:40881323"

/translation="MPCAGSRGSPAGDRLRLARLALVLLGWVSAPSSVPSSSTS
 PAAFLASGAQPPPAERCPACCEAARTKVCNRNLLLEVPADLPYVNRNLFITGNO
 MTVLPAGAFARQPPADLALNLSGNHLKEVCAGAFELPGLRLDLSHNPLTNLSAF
 APAGSNASVSPLEELILNHIVPPEDQKQSGFEGMVAFGMVAALRSGLALRGL
 TCLEASHNHFELPDLRLAQLPSRLYLDLRNNSLSVLTYSFRNLTHLESILHLEDNAL
 KVLHNSTLAEWQGLAHVKVFLDNNPWCDCYMDADMWALVLTETVVPDKARLTCAPEK
 MNRNGLDLNSDLDCDAVLPSQLSQTSYVFLGIVLALIGAIFLLVLYLNRRKGIKQWH
 NIRDACRDHMEGYHYRYEINADPRLTNLSNSDV"

ORIGIN

Alignment Scores:
 Pred. No.: 50.8 Length: 2557
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX961912 (1-2557)


```

Alignment Scores:
Pred. No.: 54.2 Length: 2714
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x AB168308 (1-2714)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
|||||
Db 1853 GCCCTGATAGGCGCTATTTTCTCTCTG 1879

RESULT 34
HSA012159 5551 bp DNA linear PRI 15-APR-2005
LOCUS Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
DEFINITION
ACCESSION AJ012159
VERSION AJ012159.1 GI:3805946
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10366710
REFERENCE 2 (bases 1 to 5551)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK
FEATURES
source
1. .5551
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
2698. .2703
/bound moiety="Sp1"
misc_binding
2704. .2709
/bound moiety="Sp1"
misc_binding
2716. .5400
gene
2716. .2800
exon
/genes="5T4"
/genes="5T4"
evidence=experimental
intron
2801. .3092
/genes="5T4"
evidence=experimental
exon
3093. .5400
/genes="5T4"
evidence=experimental
CDS
3431. .4693
/genes="5T4"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA09930.1"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR01611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCSRGPAAGDGRLRRLARLALVLLGWSSSSPTSSASSPSSS
APPLASAVSAQPLPLDQCPCSEARTVKVNRNLTEVPTDLPAYVYRNLFITGNQ
LAVLPAGAFARRPPLAEALANLSSRLDEVRAGAFELFSLRQLDLSHNPADLSPP
AFSGSNASAPSPVLVLIINHVPPEDEQRNRSFEGMVVAALLAGRALQRLRLELA
SNHFLVLPDVLQALPSLRHLDSNLSVSLTYVPSNLTSLSLHLEDNALKVLHNG
TLAEQLQLPRIRVFLDNNPVCDCMADMTWLKETEYVCGKDRLTCTAYPEKMRNVL
LELSADLDCDPILPSPSQTSYVFLGIVLALIGAIFLLVLYLNRRKGIKKWNNIRDAC
RDHMEGYHYRYEINADPRLTNLSSNSDV"
sig_peptide 3431. .3516
mat_peptide 3517. .4690
polyA_signal 5331. .5336
polyA_signal 5380. .5385
ORIGIN
Alignment Scores:
Pred. No.: 118 Length: 5551
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x HSA012159 (1-5551)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
|||||
Db 4520 GCCCTGATAGGCGCTATTTTCTCTCTG 4546

RESULT 35
MMU012160 7942 bp DNA linear ROD 15-APR-2005
LOCUS Mus musculus 5T4 oncofetal trophoblast glycoprotein gene.
DEFINITION
ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10366710
REFERENCE 2 (bases 1 to 7942)
AUTHORS Myers, K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK
FEATURES
source
1. .7942
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="Lambda Dash"
3108. .3113
/bound moiety="Sp1"
misc_binding
3114. .3119
/bound moiety="Sp1"
misc_binding
3124. .5779
gene
/genes="5T4"
3124. .3151
exon
/genes="5T4"
3152. .3450
intron
/genes="5T4"

```


APFLASAVSAQPLPDQCPCSEAAARTKVCNRLTEVPTDLPAYVRNLFRTGNO
LAVLPAGAFARRPPLAAGALNLSGRDLDEVRAGAFELPSLROLDLSHNPLADLSPP
AFSGNSASVSAPSPLVELLNIHVPPEDQRNRSFEGMVVAALLAGRALQGLRLLELA
SNFHYLPDRVLAQPSLRHLDSNLSVLTFSFRNLTHLSLHSDNALKVLHNG
TLELSADLDCDFILPSPQTSYVFLGIVLALIGAFLLVLYLNKRGKIKKMMENIRDAC
RDHMEGYHYREINADPRLTNLSSNDV"
116817..116822
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
116836
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
121909_tag="RP3-492P14.1-001"
/note="Clone_right_end: RP3-492P14"

polyA_signal

polyA_site

misc_feature

ORIGIN

Alignment Scores:
Pred. No.: 3,42e+03 Length: 121909
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-6 (1-9) x H5J492P14 (1-121909)

Qy 1 AlaLeuileGlyAlailePheLeuLeu 9

Db 112059 GCCCTGATAGCGCTATTTCCTCTG 112085

RESULT 37

AC158516/c

LOCUS

DEFINITION Mus musculus BAC clone RP24-511A23 from chromosome 9, complete

ACCESSION AC158516

VERSION AC158516.2 GI:63025421

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 167046)

Adams,S., Cotton,M. and Haglund,K.

The sequence of Mus musculus BAC clone RP24-511A23

Unpublished (2001)

2 (bases 1 to 167046)

Wilson,R.K.

Direct Submission

Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 167046)

Wilson,R.K.

Direct Submission

Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 167046)

Wilson,R.K.

Direct Submission

Submitted (21-JUN-2005) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On May 4, 2005 this sequence version replaced gi:61656412.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0511A23

Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e. phred quality
>=30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone, fosmid clone or direct clone walk sequence.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to obtain the consensus sequence. The
assembly was confirmed by restriction digest.

This finishing standard has slightly changed from the previous
Human standard. Specifically, standards for regions of low sequence
complexity (such as dinucleotide repeats and small unit tandem
repeats) have been relaxed. These regions are very prevalent in the
mouse genome, and the return on extended finishing efforts is
minimal.

If a sequence meets the criteria of the above statement, it needs
no comments or tags. If the criteria are not met, such as ambiguous
bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:

The BAC library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

This sequence is the entire insert of the clone.

FEATURES

source

1..167046

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosomes="9"

/clone="RP24-511A23"

/clone_lib="RPCT-24"

16685..16712

/note="Sequence derived from PCR product of genomic DNA"

31565..31779

/note="Unresolved simple sequence repeat."

46721..46808

/note="Unresolved simple sequence repeat."

142336..142347

/note="Sequence derived from one plasmid subclone."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

4.83e+03

40.00

100.0%

100.0%

100.0%

9

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

167046

9

0

0

0

0

0

US-10-774-176-6 (1-9) x AC158516 (1-167046)

Qy 1 AlaLeuileGlyAlailePheLeuLeu 9

Db 109731 GCTCTGATAGCGCTATTTCCTCTC 109705

RESULT 38

AC128294/c

LOCUS

DEFINITION Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.

ACCESSION AC128294

VERSION AC128294.3 GI:25083347

HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
 REFERENCE
 1 (bases 1 to 210237)
 Muzny, D., Marle, E., Metzker, M., Lee, J., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alabrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biewala, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Cesear, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M.,
 George, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,
 Hollings, B., Howells, S., Hui, Y., Hume, J., Idiebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowitz, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshewa, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 McWhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minje, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokeneme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poinexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Rellis, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shateman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.
 Direct Submission
 2 (bases 1 to 210237)
 Worley, K.C.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 210237)
 Rat Genome Sequencing Consortium.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:23265004.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Aclis

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: CH230-176H20
 Center clone name: CH230-176H20
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 201781 bases at least Q40
 Consensus quality: 203921 bases at least Q30
 Consensus quality: 205310 bases at least Q20
 Estimated insert size: 205531; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 210237: contig of 210237 bp in length.

FEATURES
 source
 1. 210237
 /location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-176H20"
 1. 1142
 /note="wgs end extension
 clone_end:77"
 2177..144799
 /note="clone boundary
 clone_end:77"
 misc_feature
 1. 1142
 /note="wgs end extension
 clone_end:77"
 2177..144799
 /note="clone boundary
 clone_end:77"
 misc_feature
 end_sequence:BH360464"
 complement(206062..206961)
 /note="clone boundary
 clone_end:Sp6"
 site:
 end_sequence:BH360465"
 208307..210237
 /note="wgs end extension
 clone_end:Sp6"
 misc_feature
 end_sequence:BH360465"
 208307..210237
 /note="wgs end extension
 clone_end:Sp6"
 misc_feature
 end_sequence:BH360465"
 208307..210237
 /note="wgs end extension
 clone_end:Sp6"

ORIGIN
 Alignment Scores:
 Pred. No.: 6.2e+03 Length: 210237
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0
 US-10-774-176-6 (1-9) x AC128294 (1-210237)
 QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9

```

Baylor Plaza, Houston, TX 77030, USA
3 (Bases 1 to 239076)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:22857070.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOPI
Center clone name: CH230-87110
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 228642 bases at least Q40
Consensus quality: 232629 bases at least Q30
Consensus quality: 234041 bases at least Q20
Estimated insert size: 231522; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 234710: contig of 234710 bp in length
234711 234810: gap of unknown length
234811 235924: contig of 1114 bp in length
235925 236024: gap of unknown length
236025 237314: contig of 1290 bp in length
237315 237414: gap of unknown length
237415 239076: contig of 1662 bp in length.
Location/Qualifiers
1. 239076
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-87110"
234711..234810
/estimated_length=unknown
235925..236024
/estimated_length=unknown
237315..237414
/estimated_length=unknown
ores:
7.13e+03 Length: 239076
40.00 Matches: 9
100.0% Conservative: 0
100.0% Mismatches: 0
100.0% Indels: 0

```

Phunthang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 158710)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collumore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McClean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-NOV-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 158710)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collumore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McClean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (21-DEC-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 21, 2004 this sequence version replaced gi:55733999.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L25036
Center clone name: 112_G_5
----- Location/Qualifiers

FEATURES

DB: 14 Gaps: 0

US-10-774-176-6 (1-9) x AC109692 (1-239076)

Qy 1 AlaleulleGlyAlaIlePheLeuLeu 9
|||||
15622 GCCTGATAGCGCGCTATTTCCTCCCTC 15596

Db 15622 GCCTGATAGCGCGCTATTTCCTCCCTC 15596

RESULT 40

LOCUS AR377516 1494 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 2522 from patent US 6605709.

ACCESSION AR377516

VERSION AR377516.1 GI:40080698

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1494)
Braton, G.L.
Nucleic acid and amino acid sequences relating to Proteus mirabilis for diagnostics and therapeutics
Patent: US 6605709-A 2522 12-AUG-2003;
Genome Therapeutics Corporation; Waltham, MA

TITLE Location/Qualifiers

JOURNAL

FEATURES

source 1..1494
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 49.2 Length: 1494
Score: 39.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 97.5% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AR377516 (1-1494)

Qy 1 AlaleulleGlyAlaIlePheLeuLeu 9
|||||
535 GCCTGATAGCGCGCTATTTCCTATTA 561

Db 535 GCCTGATAGCGCGCTATTTCCTATTA 561

RESULT 41

LOCUS AC126539 158710 bp DNA linear ROD 21-DEC-2004

DEFINITION Mus musculus chromosome 8, clone RP24-112G5, complete sequence.

ACCESSION AC126539

VERSION AC126539.9 GI:56744340

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 158710)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 8, clone RP24-112G5
Unpublished

REFERENCE 2 (bases 1 to 158710)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collumore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gorm, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., McClean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,


```

Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-MAR-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 206384)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepe,Y., Collumore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-APR-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 206384)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepe,Y., Collumore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-JUN-2005) Broad Institute of MIT and Harvard, 320
Charles Street, Cambridge, MA 02141, USA
On Jun 28, 2005 this sequence version replaced gi:62822033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
----- Project name: L31534
Center project name: L31534
Center clone name: 18_G_2
-----

```

```

FEATURES
source
Location/Qualifiers
1..206384
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone="RP23-18G2"
/clone_lib="RPCI-23 Female Mouse BAC"
complement(1..67)
/rpt_family="RSINE1"
1862..1882
/rpt_family="(T)n"
complement(3661..3723)
/rpt_family="ID_B1"
4258..4377
/rpt_family="B3"
complement(5019..5452)
/rpt_family="MLTIP1"
5468..5629
/rpt_family="(CATA)n"
5541..5693
/rpt_family="(CATATA)n"
5874..5896
/rpt_family="(CA)n"
complement(6607..6863)
/rpt_family="B4"
complement(7583..7625)
/rpt_family="L1"
7619..8425
/rpt_family="L1_MM"
complement(8477..8665)
/rpt_family="URR1B"
10367..10389
/rpt_family="AT rich"
complement(10399..10489)
/rpt_family="B3A"
10570..10661
/rpt_family="(TCCC)n"
complement(11029..11217)
/rpt_family="B2_Mm1"
11668..11726
/rpt_family="AT rich"
complement(11729..13598)
/rpt_family="L1"
13597..15299
/rpt_family="L1_MM"
14009..14035
/note="single clone coverage"
15302..15324
/rpt_family="(CAAA)n"
complement(15431..15592)
/rpt_family="RSINE1"
16179..16853
/rpt_family="Lx2B"
16851..17099
/rpt_family="Lx2B"
17106..17125
/rpt_family="(TAAAA)n"
17306..17527
/rpt_family="RMER20"
18161..18237
/rpt_family="GA-rich"
18267..18290
/rpt_family="(TG)n"
complement(18291..18509)
/rpt_family="MSTD"
complement(18654..18748)
/rpt_family="PB1D10"
18798..18819
/rpt_family="(TTTA)n"
complement(18822..18968)
/rpt_family="B1_MM"

```

```

repeat_region complement(18970. .19179)
/rpt family="B3"
repeat_region 1941. .19699
/rpt family="B4"
repeat_region 19700. .19789
/rpt family="(CA)n"
repeat_region 20392. .20436
/rpt family="CT-rich"
repeat_region 21677. .21730
/rpt family="(TG)n"
repeat_region 22703. .22740
/rpt family="(CA)n"
repeat_region 23139. .23228
/rpt family="RSINR1"
repeat_region complement(23235. .23562)
/rpt family="LI_MM"
repeat_region 23836. .23885
/rpt family="(TA)n"
repeat_region complement(24389. .24575)
/rpt family="B3A"
repeat_region 24646. .24935
/rpt family="MYS_LTR"
repeat_region 25305. .25326
/rpt family="AT-rich"
repeat_region 25819. .25941

Alignment Scores:
Pred. No.: 1.06e+04 Length: 206384
Score: 39.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 97.5% Indels: 0
DB: 9 Gaps: 0

US-10-774-176-6 (1-9) x AC158916 (1-206384)

Qy 1 AlaleulleGlyAlaIlePheluleu 9
Db 20144 GCCTTGTGGGAGCCATCTTCTATTA 20170

RESULT 43
AC107258/c AC107258 208078 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-138L4, WORKING DRAFT SEQUENCE, 3
DEFINITION unorderd pieces.
ACCESSION AC107258
VERSION AC107258.5 GI:25006628
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 208078)
AUTHORS Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,P.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,Y., Chen,Z., Chu,J.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleaveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,B., Geis,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,

```

```

Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlerbird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuewa,L., Loulaseg,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,B.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Flopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Fuazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wiecezyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 208078)
Worley,K.C.
Direct Submission
Submitted (18-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 208078)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:22857079.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOMX
Center clone name: CH230-138L4
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 196804 bases at least Q40
Consensus quality: 199513 bases at least Q30
Consensus quality: 201135 bases at least Q20
Estimated insert size: 206052; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----

```



```

Assembly program: Atlas 3.0;
Consensus quality: 180076 bases at least Q40
Consensus quality: 185112 bases at least Q30
Consensus quality: 189119 bases at least Q20
Estimated insert size: 187073; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 6 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
*
* 1 223057: contig of 223057 bp in length
* 223058 223157: gap of unknown length
* 223158 234587: contig of 11430 bp in length
* 234588 234687: gap of unknown length
* 234688 235919: contig of 1232 bp in length
* 235920 236019: gap of unknown length
* 236020 237370: contig of 1351 bp in length
* 237371 237470: gap of unknown length
* 237471 238675: contig of 1205 bp in length
* 238676 238775: gap of unknown length
* 238776 241416: contig of 2641 bp in length.
FEATURES             Location/Qualifiers
     source            1..241416
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-69E14"
     misc_feature      1..1322
                        /note="wgs end extension"
                        clone_end:Sp6"
     misc_feature      4396..4961
                        /note="clone boundary"
                        clone_end:Sp6"
                        site:ECORI
                        end_sequence:BH355423"
                        78244..79433
                        /note="wgs contig"
                        complement(150409..181071)
                        /note="clone boundary"
                        clone_end:T7"
                        site:ECORI
                        end_sequence:BH355420"
                        150661..151638
                        /note="wgs end extension"
                        clone_end:T7"
     misc_feature      169162..170914
                        /note="wgs end extension"
                        clone_end:T7"
     gap               223058..223157
                        /estimated_length=unknown
     gap               234588..234687
                        /estimated_length=unknown
     gap               235920..236019
                        /estimated_length=unknown
     gap               237371..237470
                        /estimated_length=unknown
     gap               238676..238775
                        /estimated_length=unknown
ORIGIN
Alignment Scores:
Pred. No.:           1-25e+04          Length:           241416
Score:               39.00             Matches:           8
Percent Similarity: 100.0%             Conservative:    1
Best Local Similarity: 88.9%            Mismatches:      0
Query Match:         97.5%             Indels:          0

```

```

DB:                  14              Gaps:          0
US-10-774-176-6 (1-9) x AC114100 (1-241416)
QY      1 AlaLeuileGlyAlaIlePheLeuLeu 9
        |||||:|||||
Db      149974 GCACGTGGTGGTGCCATCTTCTACTG 149948
RESULT 45
LOCUS   AX385620
DEFINITION Sequence 548 from Patent WO0214500.
ACCESSION AX385620
VERSION   AX385620.1 GI:19578750
KEYWORDS .
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE
AUTHORS   Escobedo,J., Garcia,P.D., Sudduth-Klinger,J., Reinhard,C.,
          Randazzo,F., Lamson,G., Scott,E.M., Zhang,G., Kassam,A., Pot,D. and
          Labat,I.
TITLE     Human genes and gene expression products
JOURNAL   Patent: WO 0214500-A 548 21-FEB-2002;
          CHIRON CORPORATION (US) ; Hyseq Inc. (US)
FEATURES             Location/Qualifiers
     source            1..433
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:           22.2              Length:           433
Score:               38.00             Matches:           8
Percent Similarity: 100.0%             Conservative:    1
Best Local Similarity: 88.9%            Mismatches:      0
Query Match:         95.0%             Indels:          0
DB:                  6              Gaps:          0
US-10-774-176-6 (1-9) x AX385620 (1-433)
QY      1 AlaLeuileGlyAlaIlePheLeuLeu 9
        |||||:|||||
Db      200 GCTCTTCTAGCGCCATCTTCTCTCTC 174
RESULT 46
LOCUS   AX385964/c
DEFINITION Sequence 892 from Patent WO0214500.
ACCESSION AX385964
VERSION   AX385964.1 GI:19579094
KEYWORDS .
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE
AUTHORS   Escobedo,J., Garcia,P.D., Sudduth-Klinger,J., Reinhard,C.,
          Randazzo,F., Lamson,G., Scott,E.M., Zhang,G., Kassam,A., Pot,D. and
          Labat,I.
TITLE     Human genes and gene expression products
JOURNAL   Patent: WO 0214500-A 892 21-FEB-2002;
          CHIRON CORPORATION (US) ; Hyseq Inc. (US)
FEATURES             Location/Qualifiers
     source            1..567
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
ORIGIN

```


Alignment Scores:
 Pred. No.: 29.8 Length: 567
 Score: 38.00 Matches: 8
 Percent Similarity: 100.0% Conservatives: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 95.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-6 (1-9) x AX385964 (1-567)

Qy 1 AlaLeulleGlyAlaIlePheLeu 9
 |||||:|||||:|||||:|||||
 Db 188 GCTCTCTAGGCGCATCTCTCTCTC 162

RESULT 47

AJ851373 3205 bp mRNA linear VRT 15-APR-2005
 LOCUS Gallus gallus mRNA for hypothetical protein, clone 1b1.
 DEFINITION
 AJ851373
 ACCESSION
 VERSION AJ851373.1 GI:60098352
 KEYWORDS ORP1.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE 1
 AUTHORS Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezzubov, Y., Zaim, J.,
 Fiedler, P., Kutter, S., Blagodatski, A., Kostovska, D., Kotar, M.,
 Plachy, J., Carninci, P., Hayashizaki, Y. and Buerstedde, J.M.
 TITLE Full-length cDNAs from chicken bursal lymphocytes to facilitate
 gene function analysis
 JOURNAL Genome Biol. 6 (1), R6 (2005)
 PUBMED 15642098
 REFERENCE 2 (bases 1 to 3205)
 AUTHORS Caldwell, R.B.
 TITLE Direct Submission
 JOURNAL Submitted (12-OCT-2004) Caldwell R.B., GSF - Forschungszentrum,
 Institut fuer Molekulare Strahlenbiologi, Ingolstaedter Landstr. 1,
 D-85764 Neuherberg, GERMANY

FEATURES

source
 1..3205
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="CB"
 /db_xref="taxon:9031"
 /clone="1b1"
 /cell_type="lymphocyte"
 /tissue_type="bursa"
 /clone_lib="rikeni"
 /dev_stage="2 weeks old"
 1..3205
 /locus_tag="RCJMB04_1b1"
 <1..421
 /locus_tag="RCJMB04_1b1"
 422..2287
 /locus_tag="RCJMB04_1b1"
 /note="ORF1"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAH65007.1"
 /db_xref="GI:60098353"
 /db_xref="UniProt/TREMBL:Q5P4C7"

/translation="MRLRRYSPLKVALGVALLPLFLIQRDVGDDPREPWLN
 IVDKQVLDLMLGAVENLRDMPKLIQIGAVPPEPPRSARSCPPGEVTAARPLI
 ERPQDPSGDGAKFAKQWTAESKEKRGYKHCNFAFADRIQLQALGPDNR
 PPECIDQKRCPLPTTSVIVFHNEMSLTTRVSVLHASPAALLRLRILVDDAS
 TDEYLDKDELQIVRVVQRKRLITABLLGASVSGEVLITLDAHCFCFHG
 WLEPLLRIRAEPTAVSPDITLNTFFPKVQYQKHSRGNFWSLTFGEVVP
 PRERQRKDETVPIKPTFAGGLFAISRSYFHHGSYDDQMEIWGENVENVSFRVQC
 GQQLIIPCISVGVHFSKSPHTPKGTQVLSRNQVRLAEVWMDYKEIFYRNQQA
 QNAREKTYGDIERRRLRLRNLCKNFTWILQNVPFVPLNPTSYGAILKNEGTNSC

3'UTR

ORIGIN

Alignment Scores:
 Pred. No.: 197 Length: 3205
 Score: 38.00 Matches: 7
 Percent Similarity: 100.0% Conservatives: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 95.0% Indels: 0
 DB: 5 Gaps: 0

US-10-774-176-6 (1-9) x AJ851373 (1-3205)

Qy 1 AlaLeulleGlyAlaIlePheLeu 9
 |||||:|||||:|||||:|||||
 Db 461 GCCTGTGTGGCGCATCTCTCTCTC 487

RESULT 48

AJ117351 40133 bp DNA linear BCT 30-NOV-1999
 LOCUS Zymomonas mobilis rnb operon, partial sequence; RpfN (rpfN) gene,
 complete cds; rnc operon, complete sequence; Mext (mext), putative
 regulatory protein, dihydroadipic acid synthetase (dapA),
 hypothetical ABC transporter permease protein, hypothetical ABC
 transporter ATP-binding protein, MviM (mviM), NoxX (noxX),
 molybdenum transport system protein (mod), and homocitrate
 synthase (nifv) genes, complete cds; and unknown genes.
 AJ117351
 VERSION AJ117351.1 GI:6478219
 KEYWORDS

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Zymomonas mobilis
 Zymomonas mobilis
 Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
 Sphingomonadaceae; Zymomonas.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..40133
 /organism="Zymomonas mobilis"
 /mol_type="genomic DNA"
 /strain="ZM4"
 /db_xref="taxon:542"
 /clone="fosmid 44F11"
 <1..4887
 /gene="rnb operon"
 <1..1134
 /gene="rnb operon"
 /product="16S ribosomal RNA"
 1286..1362
 /gene="rnb operon"
 /product="tRNA-Ile"
 1386..1461
 /gene="rnb operon"
 /product="tRNA-Ala"
 1741..4525
 /gene="rnb operon"
 /product="23S ribosomal RNA"
 4624..4744
 /gene="rnb operon"
 /product="5S ribosomal RNA"
 4811..4887
 /gene="rnb operon"
 /product="tRNA-Met"

gene

rRNA

tRNA

tRNA

rRNA

rRNA

tRNA

```
CDS      /note="fMet"  
complement(6131..6610)  
/note="Orf1"  
/codon_start=1  
/transl_table=11  
/product="unknown"  
/protein_id="AAF13744.1"  
/db_xref="GI:6478220"  
/translation="MPEQLRQTKRCPSDINDAENWPISPPGLGVADYVKLISV  
TREGINDWRYFARSSCRPFHEVAGRFLRIIYDKRLILDHVQGRDTSVIDSQIRLP  
HTTKRQKQAINTCPCPTMNRATANTSRIVDRILSDPMLASLAMSSETSVCWRL"  
6982..7470  
/note="Orf2"  
/codon_start=1  
/transl_table=11  
/product="unknown"  
/protein_id="AAF13745.1"  
/db_xref="GI:6478221"  
/translation="MKSQPILOTSEVSSLTPLRLFIYDDKGGYVYVWQHNVAESMPE  
IIDLRYOFKALIPDCDTLAETAPHHFAIRSAEAQNLIIPREWYFARLGLSTPL  
FDEFEAEFKLINREAAIARSEBIYCGNSQSISEIHEVAKIARQYSGVIPISVSGGG  
SV"  
8766..9770  
/note="Orf3"  
/codon_start=1  
/transl_table=11  
/product="unknown"  
/protein_id="AAF13746.1"  
/db_xref="GI:6478222"  
/translation="MILPIKARCLMVEKNRSLAILKIQVPLIOAPMSWITNADPVA  
AVSAGGIVLGPAGHYTPPKDQEIILDRVLEIRKVKSLTDKPPGLNLLLLKDDPL  
DDFTKANLAFEEGVTHFVSVGNADNRVFNLEKHGDIILHRPLTASITNNREASL  
GADVLATYDGEWGPITRAIGTFTVPTWDAVDIPYMATGGINDRGRGVKAAPFALGA  
EGVIGTFTYKENPASEITKKIVESGCCDIEFVSPKQSRITQOADIILGSLYLNK  
ENTLDDLEIRSLGGVRPAMLEKGLDEGIISVNTGIDVIRSIPTVKELVIDLLGLEEE  
"  
10310..11020  
/note="Orf4"  
/codon_start=1  
/transl_table=11  
/product="unknown"  
/protein_id="AAF13747.1"  
/db_xref="GI:6478223"  
/translation="MVDSEISTPQHVYIRQMGWLRASVLGANDGILSTSLMTGVA  
SAHGSGNILLAGMSGLIAGLAWAAGEYVSVSQHDMQADVARHEAELKANPAAEK  
HELAEIYVERGLDRQLAQVAEQLMAHNALEHLDGLDLSLARPVQALASAIIS  
FSGGAIPELTALPSPPEIINITISLISILCLAVLGMVGAHLGGANVPKALRVTFCG  
ALAMGTAAIGSFPGTQV"  
complement(11092..12216)  
/gene="rpfN"  
complement(11092..12216)  
/gene="rpfN"  
/codon_start=1  
/transl_table=11  
/product="RpfN"  
/protein_id="AAF13748.1"  
/db_xref="GI:6478224"  
/translation="MMGILDLDRLVLGKGGTIRPYVTNRHQNLANTAIGNTSSVOR  
IWGQNTHLAFTWQKLPNNRLELVAGMSGNAGAPFSSSIYCNFQNSACGNTMI  
FDSNFTYWPASMGYFKAWITPKIYFEGGAFVNPYKRANDDGFPTISKATGAM  
APFELAYTNSNDKLPRTYIGGYDGGDYDVLDAIRKYAVNTGDPFYATLNGRG  
IFPFDQMIPRPNMNSERGLTVGVAMKNVGRVAEDHFLIEIGFTQGTGFKGRDKDL  
GFMLDQMSKWTIDNIVARTSAGGSRHVPDEIMBELTYGAQISRAFRVSPNLQYI  
INPQIAPSPPTKNIKPTVGLKFTLDVLSHFNARPN"  
13492..18736  
/gene="rrnC operon"  
13492..14981  
/gene="rrnC operon"  
/product="16S ribosomal RNA"  
15133..15209  
/gene="rrnC operon"  
/product="tRNA-Ile"  
15233..15308
```

```
rrNA      /gene="rrnC operon"  
15588..18372  
/product="tRNA-Ala"  
rrNA      /gene="rrnC operon"  
18473..18593  
/product="23S ribosomal RNA"  
trNA      /gene="rrnC operon"  
18660..18736  
/product="5S ribosomal RNA"  
/gene="rrnC operon"  
/product="tRNA-Met"  
/note="fMet"  
complement(19788..20681)  
/gene="next"  
complement(19788..20681)  
/gene="next"  
/codon_start=1  
/transl_table=11  
/product="Met"  
/protein_id="AAF13749.1"  
/db_xref="GI:6478225"  
/translation="MIPTYDLNLLVSLDTLLAENSVSRAAERLGLNSAMSRILRSIR  
EIFGDPILVRAGRSVPTPRAMALHETVRHVLEGAELRMTKAPNLKQINRLFSIRA  
NEGFTFOYSGKLEIAISRAPGIRLRPAIKAEKKNMLRDQIDLDIGVIGSGPETR  
IQTIORDNFIVGVVDHPLCSLKTISVEDYAKAQHVSVSRGGLFTGPIDTELENLGLK  
RQVIAVTSFSPALSVAESTDLIASVPEQQTIEARLNMYSFPPIPVQTKAISVSMWHP  
RLDLPVHRWLREQIRKLCVI"  
20840..21457  
/note="Orf5"  
/codon_start=1  
/transl_table=11  
/product="unknown"  
/protein_id="AAF13750.1"  
/db_xref="GI:6478226"  
/translation="MNRQAIPLSRREATCCLLPATIVLLGYLTVGLPLAIPVTVH  
DYLGPMTVGMVWMASSQIATILLTQFSGFVADTRGSRVLAENVGMLLCSLSALLYAVS  
LWNGLPLKIALHLLPAGRAVLFGESLLITGVLSWISIQVGMLYAGRAMVWGISMFS  
AIGGAPLGSILLKQSGFGTVALVTAHIAITGLCLDAKNRADSSA"  
complement(22266..23021)  
/note="Orf6"  
/codon_start=1  
/transl_table=11  
/product="unknown"  
/protein_id="AAF13751.1"  
/db_xref="GI:6478227"  
/translation="MKIIFIHEAPEAGFAFEIWAHKNNYQTSYSLYOGESLPQDPK  
LDLLIILGSPQATTIECPHFNAKBSRLIADCTAAGKAVIGVCLGSLQIGESALDA  
PYTHSPATEIGKEPTITIEGKKNLKTAFPGDLGVGHWNMDMPGLTAEATIIAYSEG  
CPRIIEYTPLVYGFQCHMEFTAEVIELLIAASELSKAMSDRPFVQSPDILKQNNYD  
DMNRKLSPLDKLITRAYQEQVDVFNHSSKREKQSK"  
23354..24307  
/note="Orf7"  
/codon_start=1  
/transl_table=11  
/product="putative regulatory protein"  
/protein_id="AAF13752.1"  
/db_xref="GI:6478228"  
/translation="MQGLVAQSSRGTFVADDLPDFAPIATINPPSSRPKRIIATR  
MLNISDGPDPRLVRSSELIRAFRNLGTPALSGASYVDPRGTTIRLIALAETPDA  
RGVLSIPDEIMITHGSMGLFALASALLKPSDIAVENPGIYPAWAAFRATGAKIIGI  
PVDEGIDVEKLAIIABETPALQAVVTPPHQVPTVTTLGMGRRLLEIARQHNLT  
IEDDDYHVRFDSPKILPLVARAYHQQPIHLGPFPSKLLAPTIIRLGYVIAPPDIMARM  
MAYRESIDROGDIPLLELALASMIENGDLGRHSRKSARLCG"  
24744..25640  
/gene="dapA"  
24744..25640  
/gene="dapA"  
/codon_start=1  
/transl_table=11  
/product="dihydrodipicolinate synthetase"  
/protein_id="AAF13753.1"  
/db_xref="GI:6478229"  
/translation="MDFSGLSAPFITPNNEEGVQAAPEKILQRLQRLVDAGVNSIGVES
```

TGNYPYLTERKIQYALAVKASQIPVIAISIGAFRTDILLSEQLQKGVSAALLP
LVSYHALNEDEVFSLYETVSKHVSVPVIVDSPLISFAFSDDLQYRICQLPHIQSIK
VPSAWPNSEBAYLRQLMKSLPSSIQIGVSGDPTAANALXRGAKIWTSEWGGIFPOK

Alignment Scores:
Pred. No.: 3.09e+03 Length: 40133
Score: 38.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.0% Indels: 0
DB: 1 Gaps: 0

US-10-774-176-6 (1-9) x AF117351 (1-40133)

Qy 1 AlaleuileGlyAlaIlePheLeuLeu 9

Db 31315 GCGATTATCGGGCGGATTTCCTGCTG 31341

RESULT 49

CR522870_33/c

WPCOMMENT

Sequence split into 36 fragments LOCUS CR522870 Accession CR522870

Fragment Name	Begin	End
CR522870_01	1	110000
CR522870_02	100001	210000
CR522870_03	200001	310000
CR522870_04	300001	410000
CR522870_05	400001	510000
CR522870_06	500001	610000
CR522870_07	600001	710000
CR522870_08	700001	810000
CR522870_09	800001	910000
CR522870_10	900001	1010000
CR522870_11	1000001	1110000
CR522870_12	1100001	1210000
CR522870_13	1200001	1310000
CR522870_14	1300001	1410000
CR522870_15	1400001	1510000
CR522870_16	1500001	1610000
CR522870_17	1600001	1710000
CR522870_18	1700001	1810000
CR522870_19	1800001	1910000
CR522870_20	1900001	2010000
CR522870_21	2000001	2110000
CR522870_22	2100001	2210000
CR522870_23	2200001	2310000
CR522870_24	2300001	2410000
CR522870_25	2400001	2510000
CR522870_26	2500001	2610000
CR522870_27	2600001	2710000
CR522870_28	2700001	2810000
CR522870_29	2800001	2910000
CR522870_30	2900001	3010000
CR522870_31	3000001	3110000
CR522870_32	3100001	3210000
CR522870_33	3200001	3310000
CR522870_34	3300001	3410000
CR522870_35	3400001	3510000
CR522870_36	3500001	3610000

Continuation (34 of 36) of CR522870 from base 3300001 (CR522870 Desulfotalea psychrophil

Alignment Scores:
Pred. No.: 9.27e+03 Length: 110000
Score: 38.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.0% Indels: 0
DB: 1 Gaps: 0

US-10-774-176-6 (1-9) x CR522870_33 (1-110000)

Qy 1 AlaleuileGlyAlaIlePheLeuLeu 9

Db 31315 GCGATTATCGGGCGGATTTCCTGCTG 31341

Db 68185 GCCCTTCTGGGTGCTATTTTCTTTG 68159

RESULT 50

AE008692_18/c

WPCOMMENT

Sequence split into 21 fragments LOCUS AE008692 Accession AE008692

Fragment Name	Begin	End
AE008692_00	1	110000
AE008692_01	100001	210000
AE008692_02	200001	310000
AE008692_03	300001	410000
AE008692_04	400001	510000
AE008692_05	500001	610000
AE008692_06	600001	710000
AE008692_07	700001	810000
AE008692_08	800001	910000
AE008692_09	900001	1010000
AE008692_10	1000001	1110000
AE008692_11	1100001	1210000
AE008692_12	1200001	1310000
AE008692_13	1300001	1410000
AE008692_14	1400001	1510000
AE008692_15	1500001	1610000
AE008692_16	1600001	1710000
AE008692_17	1700001	1810000
AE008692_18	1800001	1910000
AE008692_19	1900001	2010000
AE008692_20	2000001	2100000

Continuation (19 of 21) of AE008692 from base 1800001 (AE008692 Zymomonas mobilis subsp.

Alignment Scores:

Pred. No.: 9.27e+03 Length: 110000
Score: 38.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 95.0% Indels: 0
DB: 1 Gaps: 0

US-10-774-176-6 (1-9) x AE008692_18 (1-110000)

Qy 1 AlaleuileGlyAlaIlePheLeuLeu 9

Db 85388 GCGATTATCGGGCGGATTTCCTGCTG 85362

Search completed: April 25, 2006, 20:39:52

Job time : 3079.7 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 25, 2006, 10:26:14 ; Search time 295.3 Seconds
(without alignment)
203.123 Million cell updates/sec

Title: US-10-774-176-5

Perfect score: 44

Sequence: 1 FLTCGNQLAV 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US1074176/runat 24042006 165112 19185/app query.fasta.1
-DB=N Geneseq -QFMT=fastap -SUPPLX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNIT5=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000
-LOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HRAPISE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs805p
-USER=US1074176 @CGN 1 1 3463 @runat 24042006 165112 19185 -NCPV=6 -ICPV=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAEN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	44	100.0	343 8	ABX40605 Bovine ES
2	44	100.0	505 14	ACL56146 Human col
3	44	100.0	1260 6	ABK87175 cDNA enco
4	44	100.0	1260 10	ADB97513 Feline 5T

5	44	100.0	1260	10	ADB97452	ADB97452 Human enco
6	44	100.0	1263	3	AA27058	Human 5T4
7	44	100.0	1263	4	AAF89736	Nucleotid
8	44	100.0	1263	6	ABR87174	cDNA enco
9	44	100.0	1331	8	AAJ56199	Human LRR
10	44	100.0	2020	10	ADJ56299	Human cDN
11	44	100.0	2053	8	ACC51052	Human bla
12	44	100.0	2053	8	ABX76332	Lung canc
13	44	100.0	2053	8	ABX56197	Human LRR
14	44	100.0	2053	8	AAJ56200	Human LRR
15	44	100.0	2053	11	ADN38721	Cancer/an
16	44	100.0	2053	12	ADL06473	Human tum
17	44	100.0	2053	12	ADN03961	Antipsori
18	44	100.0	2053	12	ADR25444	Breast ca
19	44	100.0	2053	13	ACN38510	Tumour-as
20	44	100.0	2053	13	ADV35098	Human cDN
21	44	100.0	2338	5	AAJ87175	DNA enco
22	44	100.0	2359	4	AAK94253	Human ful
23	44	100.0	2359	12	ADL30831	Full leng
24	44	100.0	2361	4	AAK94254	Human ful
25	44	100.0	2361	12	ADI26162	Human cDN
26	44	100.0	2361	12	ADL30833	Full leng
27	40	90.9	927	6	ABT07721	Breast ca
28	40	90.9	927	8	ABX76333	Lung canc
29	40	90.9	927	10	ADB80503	Ovarian c
30	40	90.9	927	11	ADN38723	Cancer/an
31	40	90.9	1156	6	ABV99349	Human NOV
32	38	86.4	972	10	ADC92179	E. faeciu
33	38	86.4	1281	3	AAA27059	Mouse 5T4
34	38	86.4	2557	12	ADI26160	Human cDN
35	38	86.4	2557	12	ADI26158	Human cDN
36	38	86.4	2557	12	ADO35939	Novel mou
37	37	84.1	342	11	ACH94847	Klebsiell
38	37	84.1	1553	4	ABL24406	Drosophil
39	36	81.8	303	6	ABZ32426	Candida a
40	36	81.8	332	2	AAV90012	EST clone
41	36	81.8	396	4	ABA09091	Human sec
42	36	81.8	407	3	AAJ00706	Human sec
43	36	81.8	1383	8	ACA48087	Prokaryot
44	36	81.8	2273	13	ADS11585	Human the
45	36	81.8	3109	4	AAJ06046	Human rep
46	36	81.8	3109	4	ABL98611	Human tes
47	36	81.8	3180	11	ACL28854	Rice abio
48	36	81.8	3180	12	ADI45342	Rice isop
49	36	81.8	5833	5	AAJ85053	DNA enco
50	36	81.8	5833	9	ADA02798	Human TNF
51	36	81.8	52754	10	ADB72536	Human TNF
52	36	81.8	52754	10	ADC85278	Human Tnf
53	36	81.8	52754	12	ADM74393	Human car
54	36	81.8	65274	14	ADZ12653	Human can
55	36	81.8	261817	11	ACN45182	Human gen
56	35	79.5	348	5	AAJ82724	DNA enco
57	35	79.5	353	13	ACN63013	Cotton ca
58	35	79.5	367	3	AAJ02090	Human sec
59	35	79.5	467	9	ACH18168	Human adu
60	35	79.5	503	9	ACH1084	Human foe
61	35	79.5	568	5	ADI69408	Human ova
62	35	79.5	568	5	ADI75750	Human ova
63	35	79.5	617	9	ACH04278	Human CDN
64	35	79.5	710	5	AAJ90922	DNA enco
65	35	79.5	720	5	ADL40986	Human ova
66	35	79.5	731	5	AAJ68769	DNA enco
67	35	79.5	1066	5	ADL62112	Human ova
68	35	79.5	1069	3	AAJ38321	Human tra
69	35	79.5	1071	3	AAJ36234	cDNA enco
70	35	79.5	1076	4	AAJ22744	Human cDN
71	35	79.5	1082	12	ADQ95961	T cell ac
72	35	79.5	1109	10	ADC77522	Human prv
73	35	79.5	1109	12	ADN02630	Liver dis
74	35	79.5	1109	13	ADP24841	PRO polyP
75	35	79.5	1109	14	ADY17292	DNA enco
76	35	79.5	1109	14	ADY20348	DNA enco
77	35	79.5	1113	4	AAD08383	Human sec

78	35	79.5	1155	4	AAS22508	Aas22508 Human cDN	151	34	77.3	569	4	AA118835	Aa118835 Probe #87
79	35	79.5	1174	5	AAS2728	Aas2728 DNA encod	152	34	77.3	569	4	ABA31014	AbA31014 Probe #94
80	35	79.5	1218	13	AD743431	Adt43431 Bacterial	153	34	77.3	569	4	AAK38063	AaK38063 Human bon
81	35	79.5	1413	11	ADM92579	Adm92579 SNP-conta	154	34	77.3	569	4	AAK12344	AaK12344 Human bra
82	35	79.5	1659	8	ACA52837	Aca52837 Prokaryot	155	34	77.3	569	4	ABS37682	AbS37682 Human liv
83	35	79.5	13401	4	AAK89558	Aak89558 Human dig	156	34	77.3	569	6	ABS12070	AbS12070 Human gen
84	35	79.5	13403	4	AAK89557	Aak89557 Human dig	c 157	34	77.3	577	4	ABA63915	AbA63915 Human foe
85	35	79.5	23378	4	ABK42627	Abk42627 Genomic s	c 158	34	77.3	577	4	AAI44035	AaI44035 Probe #12
86	35	79.5	23378	4	AAK89555	Aak89555 Human dig	c 159	34	77.3	577	4	ABS37754	AbS37754 Human liv
87	35	79.5	23378	4	AAK89555	Aak89555 Human dig	c 160	34	77.3	577	6	ABS12158	AbS12158 Human gen
88	35	79.5	27082	4	AAK70473	Aak70473 Connectiv	c 161	34	77.3	760	6	AAAS17960	AaA17960 P. patens
89	35	79.5	83698	6	ABN85767	Abn85767 Arabidops	162	34	77.3	798	4	AAH76199	AaH76199 Human dru
90	35	79.5	110000	6	ABQ69245	Abq69245 Arabidops	163	34	77.3	858	10	ADC39084	AdC39084 Novel hum
91	35	79.5	110000	13	ABD32966_08	Abd32966 (9 of	164	34	77.3	870	6	ABK75328	AbK75328 Nucleotid
92	35	79.5	139389	6	ABK84795	Abk84795 Human cDN	165	34	77.3	891	6	ABL59146	AbL59146 Bacillus
93	35	79.5	139389	9	ADB70369	Adb70369 PAC 6802	c 166	34	77.3	891	8	ACA53506	Aca53506 Prokaryot
94	35	79.5	139389	12	ADJ37148	Adj37148 Human mal	c 167	34	77.3	1127	7	ADR19696	AdR19696 Human dru
95	35	79.5	147620	10	ADL13739	Adl13739 Osteoarth	c 168	34	77.3	1242	8	ACA51261	Aca51261 Prokaryot
96	35	79.5	147620	12	ADQ19948	Adq19948 Human sof	c 169	34	77.3	1242	8	ACA52180	Aca52180 Prokaryot
97	35	79.5	169865	12	ADQ97056	Adq97056 Human can	c 170	34	77.3	1352	8	ACA48629	Aca48629 Prokaryot
98	35	79.5	213251	6	ABQ67193	Abq67193 Listeria	171	34	77.3	1377	13	ADX11587	AdX11587 Plant ful
99	35	79.5	245531	13	ABD33022	Abd33022 Human can	172	34	77.3	1526	13	ADX14741	AdX14741 Plant ful
100	34	77.3	177	4	AAI23624	Aai23624 Probe #13	173	34	77.3	1573	13	ADX31327	AdX31327 Plant ful
101	34	77.3	177	4	AAI21686	Aai21686 Probe #11	174	34	77.3	1746	10	ADB25534	AdB25534 Brassica
102	34	77.3	177	4	ABA76137	AbA76137 Human foe	175	34	77.3	1764	8	ACA47392	Aca47392 Prokaryot
103	34	77.3	177	4	ABA66764	AbA66764 Human foe	176	34	77.3	1797	10	AAD57635	Aad57635 Rice dise
104	34	77.3	177	4	ABA68733	AbA68733 Human foe	177	34	77.3	1845	13	ADX14584	AdX14584 Plant ful
105	34	77.3	177	4	AAI46975	Aai46975 Probe #15	178	34	77.3	1874	13	ADS48405	AdS48405 Bacterial
106	34	77.3	177	4	AAI56794	Aai56794 Probe #25	179	34	77.3	1944	8	ABX08292	AbX08292 Moss lipi
107	34	77.3	177	4	AAI48935	Aai48935 Probe #17	180	34	77.3	1962	8	ABX08291	AbX08291 Moss lipi
108	34	77.3	177	4	ABA48851	AbA48851 Human bre	181	34	77.3	1978	2	AAK32597	AaK32597 RANK-1 ge
109	34	77.3	177	4	ABA50766	AbA50766 Human bre	182	34	77.3	1978	8	ABX95248	AbX95248 Rice anky
110	34	77.3	177	4	ABA33828	AbA33828 Probe #12	183	34	77.3	1978	10	ADD35917	AdD35917 Rice part
111	34	77.3	177	4	ABA35697	AbA35697 Probe #14	184	34	77.3	1978	10	ADD35920	AdD35920 Rice part
112	34	77.3	177	4	AAK40920	Aak40920 Human bon	185	34	77.3	2069	13	ADR07772	AdR07772 Full leng
113	34	77.3	177	4	AAK42859	Aak42859 Human bon	c 186	34	77.3	2083	8	ACC46090	Acc46090 Human dit
114	34	77.3	177	4	AAK51595	Aak51595 Human bon	187	34	77.3	2090	13	ADX30345	AdX30345 Plant ful
115	34	77.3	177	4	AAK17076	Aak17076 Human bra	188	34	77.3	2139	6	AAAS17966	AaA17966 P. patens
116	34	77.3	177	4	ABS40501	AbS40501 Human liv	189	34	77.3	2172	13	ADX11901	AdX11901 Plant ful
117	34	77.3	177	4	ABS50394	AbS50394 Human liv	190	34	77.3	2175	4	ABL08537	AbL08537 Drosophill
118	34	77.3	177	4	ABS42490	AbS42490 Human liv	c 191	34	77.3	2202	8	ACA37237	Aca37237 Prokaryot
119	34	77.3	177	5	AAI07376	Aai07376 Probe #73	192	34	77.3	4807	4	ABL08536	AbL08536 Drosophill
120	34	77.3	177	5	AAI09239	Aai09239 Probe #92	193	34	77.3	5000	4	AAAS14506	AaA14506 Human GGT
121	34	77.3	177	6	ABS16914	AbS16914 Human gen	c 194	34	77.3	5365	13	ADI67007	AdI67007 Novel Lac
122	34	77.3	177	6	ABS14876	AbS14876 Human gen	c 195	34	77.3	8757	12	ADOS9875	AdO59875 Novel car
123	34	77.3	186	4	AAI27771	Aai27771 Probe #17	c 196	34	77.3	35058	4	ABL05556	AbL05556 Drosophill
124	34	77.3	186	4	ABA40638	AbA40638 Probe #19	c 197	34	77.3	40304	9	ADA03014	Ada03014 Human NCF
125	34	77.3	186	4	ABA50757	AbA50757 Human bon	c 198	34	77.3	40304	10	ADB72752	AdB72752 Human NCF
126	34	77.3	186	4	AAK24758	Aak24758 Human bra	c 199	34	77.3	40304	10	ADC85494	Adc85494 Human Ncf
127	34	77.3	186	4	AAK24758	Aak24758 Human bra	c 200	34	77.3	40304	12	ADM74609	AdM74609 Human car
128	34	77.3	186	4	ABS50348	AbS50348 Human liv	201	34	77.3	54642	14	ABBS39163	AbB39163 L. pneumo
129	34	77.3	186	6	ABS24232	AbS24232 Human gen	202	34	77.3	76798	6	ABN97454	AbN97454 Gene #395
130	34	77.3	408	4	AAI14422	Aai14422 Probe #43	203	34	77.3	80332	11	ACN44842	AcN44842 Cyclin-de
131	34	77.3	408	4	ABA56147	AbA56147 Human foe	204	34	77.3	86574	13	ABK83560	AbK83560 Human cDN
132	34	77.3	408	4	AAI35795	Aai35795 Probe #44	c 205	34	77.3	86574	6	ABK83560	AbK83560 Human cDN
133	34	77.3	408	4	ABA5644	AbA5644 Human bre	c 206	34	77.3	86574	13	ADR52822	AdR52822 Drug ther
134	34	77.3	408	4	ABA25798	AbA25798 Probe #42	c 207	34	77.3	98606	11	ACN43868	AcN43868 Mouse gen
135	34	77.3	408	4	AAK29831	Aak29831 Human bon	c 208	34	77.3	110000	4	AAI99682	Aai99682 (18 o
136	34	77.3	408	4	AAK04337	Aak04337 Human bra	c 209	34	77.3	110000	4	AAI99683	Aai99683 (17 o
137	34	77.3	408	4	ABS29479	AbS29479 Human liv	210	34	77.3	110000	10	ADF77343	AdF77343 (11 o
138	34	77.3	408	5	AAI04245	Aai04245 Probe #42	211	34	77.3	110000	10	ADH10017	AdH10017 Human chr
139	34	77.3	408	6	ABS04392	AbS04392 Human gen	212	34	77.3	110000	12	ADQ97266	AdQ97266 Human can
140	34	77.3	428	12	ADL10337	Adl10337 Cat flea	213	34	77.3	110000	14	ABBS39174	AbB39174 (4 of
141	34	77.3	475	4	AAI12495	Aai12495 Probe #24	214	34	77.3	110000	14	ABBS39174	AbB39174 (5 of
142	34	77.3	475	4	ABA54200	AbA54200 Human foe	c 215	34	77.3	110000	14	ABBS39175	AbB39175 (15 of
143	34	77.3	475	4	AAI33850	Aai33850 Probe #25	c 216	34	77.3	110000	14	ABBS39175	AbB39175 (16 of
144	34	77.3	475	4	ABA33747	AbA33747 Human bre	c 217	34	77.3	110000	14	ABBS39175	AbB39175 (17 of
145	34	77.3	475	4	AAK23951	Aak23951 Probe #24	218	34	77.3	110000	14	ABBS39175	AbB39175 (18 of
146	34	77.3	475	4	AAK27916	Aak27916 Human bon	219	34	77.3	155225	12	ADQ59197	AdQ59197 MSI-H car
147	34	77.3	475	4	AAK02477	Aak02477 Human bra	220	34	77.3	172637	6	ABN83124	AbN83124 Human vol
148	34	77.3	475	4	ABS27499	AbS27499 Human liv	221	34	77.3	172637	14	ABE80195	AbE80195 Human tra
149	34	77.3	475	5	AAI02405	Aai02405 Probe #23	c 222	34	77.3	231004	12	ADQ97855	AdQ97855 Mouse can
150	34	77.3	475	6	ABS02371	AbS02371 Human gen	223	34	77.3	237961	6	ABQ80552	AbQ80552 Human Can

224	34	77.3	300000	10	ADBE6352	AdB86352 Human PTP	297	33	75.0	1431	5	ABV24367	Abv24367 Human pro
225	34	77.3	300001	12	AD014076	Ad014076 Human pro	298	33	75.0	1431	5	ABV23854	Abv23854 Human pro
c 226	34	77.3	317425	14	ABE35720	Aeb35720 L. pneumo	299	33	75.0	1431	5	ABV29731	Abv29731 Human pro
227	34	77.3	335913	5	AA161371	AA161371 Soybean 2	c 300	33	75.0	1434	8	ACA43662	AcA43662 Prokaryot
228	34	77.3	335913	5	AA161372	AA161372 Soybean 2	c 301	33	75.0	1579	8	ACA93432	AcA93432 Human sec
229	34	77.3	349980	5	AAH41224	AAH41224 Pyrococcus	c 302	33	75.0	1614	13	ADX11499	Adx11499 Plant ful
c 230	33	75.0	159	6	ABK79888	Abk79888 Bacillus	c 303	33	75.0	1647	10	ADB80160	AdB80160 Mycobacte
c 231	33	75.0	232	14	ABEB9467	Aeb89467 Isolated	304	33	75.0	1680	6	ABT12982	Abt12982 Arabidops
c 232	33	75.0	249	8	ACA35672	AcA35672 Prokaryot	305	33	75.0	1686	2	ABT29716	Abt29716 Phosphati
233	33	75.0	278	10	ABX83960	ABX83960 Corn ear-	306	33	75.0	1707	9	ADB10219	AdB10219 Allostoc
234	33	75.0	317	5	AAV79371	AAV79371 Eucalyptu	307	33	75.0	1707	9	ADB10223	AdB10223 Allostoc
235	33	75.0	321	5	ABV06804	Abv06804 Human pro	308	33	75.0	1707	9	ADB10221	AdB10221 Allostoc
236	33	75.0	324	11	ACH94472	Ach94472 Klebsiell	c 309	33	75.0	2104	10	ADA54012	Ada54012 Human cod
237	33	75.0	334	6	ABL86842	AbL86842 Human ova	310	33	75.0	2163	8	ACA24457	AcA24457 Prokaryot
238	33	75.0	400	5	ABV36750	Abv36750 Human pro	311	33	75.0	2280	6	ABZ78215	Abz78215 A. niger
c 239	33	75.0	433	2	AAV90130	AAV90130 EST clone	312	33	75.0	2311	12	ADQ64722	AdQ64722 Novel hum
240	33	75.0	458	4	AAI02538	AAI02538 Human rep	c 313	33	75.0	2349	11	ABD05373	AbD05373 Pseudomon
241	33	75.0	504	9	ACH36471	Ach36471 Human end	c 314	33	75.0	2535	2	AAAX20674	Aax20674 Polynucle
242	33	75.0	546	5	ABV06466	Abv06466 Human pro	c 315	33	75.0	2756	12	ADJ39544	Adj39544 Plant cDN
243	33	75.0	570	4	AAH56947	Aah56947 P. patens	c 316	33	75.0	2976	11	ACL26858	AcL26858 Rice abio
244	33	75.0	570	5	AAH50940	Aah50940 Lipid deg	317	33	75.0	3657	5	ABV25797	Abv25797 Human pro
245	33	75.0	588	5	ABV54666	Abv54666 Human pro	c 318	33	75.0	4179	8	ACA41579	AcA41579 Prokaryot
246	33	75.0	601	5	ABV36421	Abv36421 Human pro	c 319	33	75.0	4182	10	ABZ41670	Abz41670 N. gonorr
247	33	75.0	601	5	ABV45440	Abv45440 Human pro	320	33	75.0	4182	10	ABZ41188	Abz41188 N. gonorr
c 248	33	75.0	654	5	AA594238	AA594238 DNA encod	c 321	33	75.0	5023	2	AAQ13821	Aaq13821 Hamster G
c 249	33	75.0	669	5	ABV15635	Abv15635 Human pro	322	33	75.0	5627	4	AAI99020	Aai99020 Human exc
c 250	33	75.0	724	6	ABL01444	AbL01444 Murine ap	323	33	75.0	5627	5	AAI63370	Aai63370 Human kid
251	33	75.0	735	4	AAI07288	AAI07288 Human rep	c 324	33	75.0	7277	8	ACA41347	AcA41347 Prokaryot
252	33	75.0	771	6	ABN89008	Abn89008 Human pro	325	33	75.0	10225	4	AAI99021	Aai99021 Human exc
c 253	33	75.0	792	6	ABQ42606	Abq42606 Oligonuc	326	33	75.0	10225	4	AAI99022	Aai99022 Human exc
254	33	75.0	792	6	ABQ42607	Abq42607 Oligonuc	327	33	75.0	10225	5	AAI63371	Aai63371 Human kid
c 255	33	75.0	795	6	ABQ42584	Abq42584 Oligonuc	328	33	75.0	10225	5	AAI63372	Aai63372 Human kid
256	33	75.0	795	6	ABQ42585	Abq42585 Oligonuc	c 329	33	75.0	17782	3	AAA81530	Aaa81530 N. mening
257	33	75.0	800	2	AAV27566	AAV27566 Human glu	c 330	33	75.0	43680	6	ABK62024	Abk62024 Human gen
c 258	33	75.0	800	3	AAAC40687	AAc40687 Arabidops	c 331	33	75.0	56018	12	ADQ97215	AdQ97215 Human can
c 259	33	75.0	816	3	AAAC34805	AAc34805 Arabidops	c 332	33	75.0	65558	14	AEA61193	Aea61193 Human BCH
c 260	33	75.0	823	3	AAAC50032	AAc50032 Arabidops	c 333	33	75.0	74279	14	AEA61106	Aea61106 Human SPI
261	33	75.0	878	4	AAI92579	AAI92579 Human pol	c 334	33	75.0	110000	3	AAA81490_01	AAA81490_01
262	33	75.0	924	3	AAAC75120	AAc75120 Human ORF	c 335	33	75.0	110000	4	AAI99682_30	AAI99682_30
c 263	33	75.0	963	6	ABZ12263	Abz12263 Arabidops	c 336	33	75.0	110000	4	AAI99682_31	AAI99682_31
264	33	75.0	984	8	ACA35557	AcA35557 Prokaryot	c 337	33	75.0	110000	4	AAI99683_30	AAI99683_30
265	33	75.0	987	8	ACA32645	AcA32645 Prokaryot	c 338	33	75.0	110000	4	AAI99683_31	AAI99683_31
266	33	75.0	987	8	ACA48994	AcA48994 Prokaryot	c 339	33	75.0	110000	9	ADB12064_12	ADB12064_12
267	33	75.0	987	8	ACA51460	AcA51460 Prokaryot	c 340	33	75.0	110000	14	AEA61163_0	AEA61163_0
c 268	33	75.0	989	5	AAAD06657	AAa06657 A. thalia	c 341	33	75.0	110000	14	AEA61163_0	AEA61163_0
c 269	33	75.0	989	6	ABK65305	Abk65305 Arabidops	c 342	33	75.0	349980	3	AAE24241_04	AAE24241_04
c 270	33	75.0	989	9	ADA15512	Ada15512 DNA encod	343	32	72.7	65	6	ABN28273	AbN28273 Rat eplic
c 271	33	75.0	989	10	ADD55759	Add55759 Thalecres	c 344	32	72.7	100	8	ACD79701	AcD79701 E. coli K
c 272	33	75.0	989	10	ADD31020	Add31020 Plant yie	c 345	32	72.7	201	13	ADS41213	AdS41213 Human aut
c 273	33	75.0	989	12	ADI43800	Adi43800 Plant tra	c 346	32	72.7	264	10	ACF70724	AcF70724 Phototrab
c 274	33	75.0	989	12	ADO01730	Ado01730 Thalecres	c 347	32	72.7	264	10	ACF70727	AcF70727 Phototrab
c 275	33	75.0	989	14	AEA27070	Aea27070 Stress to	348	32	72.7	328	14	ACL62409	AcL62409 Human col
276	33	75.0	1005	11	ACH96404	Ach96404 Klebsiell	c 349	32	72.7	354	8	ACA00343	AcA00343 C. glutam
277	33	75.0	1026	12	ADH45505	Adh45505 Human mol	c 350	32	72.7	405	11	ACH98626	AcH98626 Klebsiell
278	33	75.0	1116	6	ABZ78272	Abz78272 A. niger	c 351	32	72.7	414	13	ADS51001	AdS51001 Bacterial
c 279	33	75.0	1122	3	AAAC37662	AAc37662 Arabidops	c 352	32	72.7	437	10	ADB57641	AdB57641 Toxicity-
280	33	75.0	1125	11	ACH97334	Ach97334 Klebsiell	c 353	32	72.7	437	10	ADB52147	AdB52147 Primary r
281	33	75.0	1150	12	ADJ67040	Adj67040 Human sec	c 354	32	72.7	448	6	AAI53481	Aai53481 Zinc tran
282	33	75.0	1163	6	ABN89009	Abn89009 Human pro	355	32	72.7	453	4	AAK60778	Aak60778 Human imm
c 283	33	75.0	1185	13	ADX30211	AdX30211 Plant ful	356	32	72.7	465	9	ACH27076	AcH27076 Human adu
284	33	75.0	1215	8	ACA52456	AcA52456 Prokaryot	357	32	72.7	500	14	ADZ81366	AdZ81366 Human chr
c 285	33	75.0	1243	6	ABL01445	AbL01445 Murine ap	358	32	72.7	515	4	AAK88324	Aak88324 Human dig
286	33	75.0	1279	12	ADJ75070	Adj75070 Marker ge	359	32	72.7	515	4	AAI57572	Aai57572 Human col
287	33	75.0	1281	13	ACN42532	AcN42532 Human dia	360	32	72.7	515	6	ABS99749	AbS99749 cDNA enco
c 288	33	75.0	1284	11	ACH98647	Ach98647 Klebsiell	361	32	72.7	515	10	ABH92829	AbH92829 Human col
c 289	33	75.0	1301	3	AAAC47385	AAc47385 Arabidops	c 362	32	72.7	519	5	AAH67499	Aah67499 C glutami
290	33	75.0	1317	12	ADP86420	Adp86420 Human glu	363	32	72.7	582	8	ACA37845	AcA37845 Prokaryot
291	33	75.0	1335	11	ABD05271	Abd05271 Pseudomon	364	32	72.7	586	12	ADQ18367	AdQ18367 Human sof
c 292	33	75.0	1336	3	AAAC50210	AAc50210 Arabidops	c 365	32	72.7	592	6	ABN65195	Abn65195 Human can
c 293	33	75.0	1339	3	AAAC38462	AAc38462 Arabidops	c 366	32	72.7	606	10	ADE54066	AdE54066 Human pro
294	33	75.0	1383	11	ADI31215	Adi31215 Human cDN	c 367	32	72.7	631	3	AAF12797	Aaf12797 Aspergill
295	33	75.0	1383	13	ADS83282	AdS83282 Human lym	c 368	32	72.7	631	13	ADU56838	AdU56838 Aspergill
296	33	75.0	1421	13	ACN42531	AcN42531 Human dia	c 369	32	72.7	631	14	ADZ94841	AdZ94841 Aspergill

c 370	32	72.7	652	13	ADQ54545	Adq54545 Novel can	443	32	72.7	2982	12	ADO01537	Ado01537 Human cyc
371	32	72.7	658	10	ADD34732	Add34732 Mouse mit	444	32	72.7	3034	4	ABL10104	Ab110104 Drosophil
372	32	72.7	667	6	ABO59752	Abq59752 Human col	445	32	72.7	3051	5	AAS86808	Aas86808 DNA encod
373	32	72.7	681	10	ADH84102	Adh84102 Enterococ	446	32	72.7	3147	13	ADS73663	Ads73663 M. grisea
374	32	72.7	699	6	ABN92146	Abn92146 Staphyloc	447	32	72.7	3309	5	AAS89902	Aas89902 DNA encod
375	32	72.7	699	13	ADS01841	Ads01841 Staphyloc	448	32	72.7	3385	4	ABL06022	Ab106022 Drosophil
c 376	32	72.7	712	3	RAC33120	Rac33120 Arabidops	c 449	32	72.7	3411	8	ACA35921	Aca35921 Prokaryot
c 377	32	72.7	724	3	AAA79546	Aaa79546 Pinus rad	c 450	32	72.7	3417	11	ABD00628	Abd00628 Klebsiell
c 378	32	72.7	756	5	AAR66084	Aar66084 C. glutami	c 451	32	72.7	3523	4	AAS55048	Aas55048 S. epider
c 379	32	72.7	770	4	AAL21878	Aal21878 Human bre	c 452	32	72.7	3785	4	AAK90202	Aak90202 Human dig
c 380	32	72.7	772	4	AAI97332	Aai97332 Human neu	c 453	32	72.7	3785	4	AAI57643	Aai57643 Human col
c 381	32	72.7	792	10	ADD34731	Add34731 Mouse mit	c 454	32	72.7	3785	6	ABS99820	Abs99820 Genomic D
c 382	32	72.7	793	6	ABK63603	Abk63603 Rat seque	c 455	32	72.7	3785	10	ADB92973	Adb92973 Human col
383	32	72.7	893	10	ADB59216	Adb59216 Toxicity-	c 456	32	72.7	3889	4	ABL19702	Ab119702 Drosophil
384	32	72.7	893	10	ADS53907	Ads53907 Primary r	c 457	32	72.7	3896	4	ABL07440	Ab107440 Drosophil
385	32	72.7	893	13	ADV41803	Adv41803 Rat cardi	c 458	32	72.7	4213	13	ADR06720	Adr06720 Full leng
c 386	32	72.7	1056	5	AAR66082	Aar66082 C. glutami	c 459	32	72.7	4402	5	AAS89717	Aas89717 DNA encod
387	32	72.7	1074	6	ABL91237	Ab191237 Chlamydia	c 460	32	72.7	4416	10	ACF69599	Acf69599 Photorhab
388	32	72.7	1074	6	ABL91237	Ab191237 Chlamydia	c 461	32	72.7	4532	10	ACC49350	Acc49350 Human NET
c 389	32	72.7	1105	12	ADQ22883	Adq22883 Human sof	c 462	32	72.7	4857	4	ABK43042	Abk43042 Genomic s
c 390	32	72.7	1126	2	RAX85000	Rax85000 Human sec	c 463	32	72.7	4857	9	ADB61198	Adb61198 Connectiv
c 391	32	72.7	1126	8	ACD18926	Acd18926 Novel hum	c 464	32	72.7	5055	2	AAV17236	Aav17236 DNA from
c 392	32	72.7	1126	12	ADG78317	Adg78317 Human sec	c 465	32	72.7	5106	12	ADG39768	Adg39768 Goose par
c 393	32	72.7	1126	12	ADN60608	Adn60608 Human sec	c 466	32	72.7	5278	4	ABL27914	Ab127914 Drosophil
c 394	32	72.7	1206	10	ACF67627	Acf67627 Photorhab	c 467	32	72.7	5323	6	ABV77898	Abv77898 Hypoxia-r
c 395	32	72.7	1207	6	ABQ98638	Abq98638 Human ORF	c 468	32	72.7	5323	6	ABN95133	Abn95133 Gene #163
c 396	32	72.7	1347	4	RAS23419	Ras23419 Candida a	c 469	32	72.7	5323	10	ADE84920	Ade84920 Farnesyl
c 397	32	72.7	1347	6	ABZ31751	Abz31751 Candida a	c 470	32	72.7	5324	12	ADQ86198	Adq86198 Human tum
c 398	32	72.7	1373	3	RAC76004	Rac76004 Human ORF	c 471	32	72.7	5324	12	ADQ87349	Adq87349 Human tum
c 399	32	72.7	1386	4	AAA91315	Aaa91315 NS1 prote	c 472	32	72.7	5935	5	AAD10306	Aad10306 Human Par
c 400	32	72.7	1386	6	RAD36286	Rad36286 Goose par	c 473	32	72.7	6911	13	ADR84267	Adr84267 Aspergill
c 401	32	72.7	1386	6	RAD46143	Rad46143 Goose par	c 474	32	72.7	7327	10	ADC24716	Adc24716 Human HNL
c 402	32	72.7	1386	6	RAD46143	Rad46143 Goose par	c 475	32	72.7	7848	13	ADR84458	Adr84458 Aspergill
c 403	32	72.7	1386	8	ACC69250	Acc69250 Goose par	c 476	32	72.7	7964	6	ABS98833	Abs98833 Enterococ
c 404	32	72.7	1386	10	ADI40291	Adi40291 Goose par	c 477	32	72.7	7964	6	ABS98833	Abs98833 Enterococ
c 405	32	72.7	1386	10	ABX96674	Abx96674 NS1 DNA s	c 478	32	72.7	8912	4	AAK78524	Aak78524 Human imm
c 406	32	72.7	1386	10	ABX96529	Abx96529 DNA encod	c 479	32	72.7	9857	4	AAK78525	Aak78525 Human imm
c 407	32	72.7	1444	13	ADX54965	Adx54965 Plant ful	c 480	32	72.7	9857	4	AAK78525	Aak78525 Human rep
c 408	32	72.7	1451	4	ABL19703	Ab119703 Drosophil	c 481	32	72.7	9857	4	ABL98332	Ab198332 Human tes
c 409	32	72.7	1495	4	RAS96074	Ras96074 C. glutam	c 482	32	72.7	9858	4	AAK78523	Aak78523 Human imm
c 410	32	72.7	1503	3	RAC77783	Rac77783 Human can	c 483	32	72.7	9858	4	AAK78523	Aak78523 Human tes
c 411	32	72.7	1561	4	RAS23809	Ras23809 Candida a	c 484	32	72.7	9858	4	ABL98331	Ab198331 Human tes
c 412	32	72.7	1620	6	ABS55207	Abs55207 C. glutam	c 485	32	72.7	13856	2	AAV74342	Aav74342 Staphyloc
c 413	32	72.7	1638	12	ADM31188	Adm31188 Human DNA	c 486	32	72.7	15002	6	ABQ93534	Abq93534 Human Dia
c 414	32	72.7	1638	12	ADM90770	Adm90770 Human DNA	c 487	32	72.7	17135	10	AAI43384	Aai43384 pgamma6.0
c 415	32	72.7	1731	12	ADP84823	Adp84823 Cellbioh	c 488	32	72.7	23626	4	ABL12942	Ab112942 Drosophil
c 416	32	72.7	1834	5	ABV25036	Abv25036 Human pro	c 489	32	72.7	23626	4	AAS28506	Aas28506 Genomic s
c 417	32	72.7	1834	5	ABV25570	Abv25570 Human pro	c 490	32	72.7	23626	5	ABA21480	Abas21480 Human ner
c 418	32	72.7	1834	5	ABV25416	Abv25416 Human pro	c 491	32	72.7	23626	5	ABA21480	Abas21480 Human ner
c 419	32	72.7	1834	5	ABV25061	Abv25061 Human pro	c 492	32	72.7	23626	10	ADB33334	Adb33334 Human nov
c 420	32	72.7	1877	10	AAI54313	Aai54313 Human SEC	c 493	32	72.7	23626	10	ADG41702	Adg41702 Human res
c 421	32	72.7	1884	4	AAA91313	Aaa91313 Rep prote	c 494	32	72.7	23626	11	ADI97476	Adi97476 Human res
c 422	32	72.7	1884	6	RAD36284	Rad36284 Goose par	c 495	32	72.7	23632	4	AAS28507	Aas28507 Genomic s
c 423	32	72.7	1884	6	RAD46141	Rad46141 Goose par	c 496	32	72.7	23632	5	ABA21481	Abas21481 Human ner
c 424	32	72.7	1884	6	RAD44603	Rad44603 Goose par	c 497	32	72.7	23632	5	AAS29998	Aas29998 Human lun
c 425	32	72.7	1884	8	ACC69248	Acc69248 Goose par	c 498	32	72.7	23632	10	ADB33335	Adb33335 Human nov
c 426	32	72.7	1884	10	ADI40287	Adi40287 Goose par	c 499	32	72.7	23632	10	ADG41703	Adg41703 Human res
c 427	32	72.7	1884	10	ABX96672	Abx96672 Rep DNA s	c 500	32	72.7	23632	11	ADI97477	Adi97477 Human res
c 428	32	72.7	1884	10	ABX96527	Abx96527 DNA encod	c 501	32	72.7	27811	11	ABX77186	Abx77186 Genomic D
c 429	32	72.7	1899	11	ACL28477	Acl28477 Rice abio	c 502	32	72.7	32176	4	AAI05628	Aai05628 Human rep
c 430	32	72.7	1908	13	ADS57803	Ads57803 Bacterial	c 503	32	72.7	32250	4	AAI05627	Aai05627 Human rep
c 431	32	72.7	2058	1	AAI71053	Aai71053 Escherich	c 504	32	72.7	32250	4	AAI05627	Aai05627 Human rep
c 432	32	72.7	2067	8	ABQ80242	Abq80242 Chromosom	c 505	32	72.7	37091	4	ABL14244	Ab114244 Drosophil
c 433	32	72.7	2133	13	ADU07863	Adu07863 DNA seque	c 506	32	72.7	39119	8	ABZ74034	Abz74034 Secreted
c 434	32	72.7	2160	13	ADU07895	Adu07895 DNA seque	c 507	32	72.7	39119	8	ADA98641	Ada98641 Human sec
c 435	32	72.7	2167	5	RAC82727	Rac82727 Beta vulg	c 508	32	72.7	39119	10	ADC20764	Adc20764 Human sec
c 436	32	72.7	2355	13	ADS49252	Ads49252 Bacterial	c 509	32	72.7	39119	10	ABZ67621	Abz67621 Human sec
c 437	32	72.7	2388	13	ADS73662	Ads73662 M. grisea	c 510	32	72.7	41637	9	ADA02837	Ada02837 Mouse Map
c 438	32	72.7	2445	10	ADC92229	Adc92229 E. faeciu	c 511	32	72.7	41637	10	ADB72575	Adb72575 Mouse Map
c 439	32	72.7	2487	8	ACF74161	Acf74161 Staphyloc	c 512	32	72.7	41637	12	ADC85316	Adc85316 Human Mef
c 440	32	72.7	2658	2	RAT92420	Rat92420 Cosmid cl	c 513	32	72.7	41637	12	ADM74432	Adm74432 Murine ca
c 441	32	72.7	2745	4	ABL27608	Ab127608 Drosophil	c 514	32	72.7	45774	12	ADJ12457	Adj12457 DNA fragm
c 442	32	72.7	2820	4	ABL27915	Ab127915 Drosophil	c 515	32	72.7	64275	14	AEA61168	Aea61168 Human HSD

c 516	32	72.7	75899	6	ABK85261	Abk85261 Human gen	589	31	70.5	208	8	ACA16590	ACA16590 Prokaryot
c 517	32	72.7	75899	12	AD113990	Ad113990 Human pro	590	31	70.5	208	8	ACA16870	ACA16870 Prokaryot
c 518	32	72.7	75899	14	ADZ56505	Adz56505 Human pro	591	31	70.5	208	8	ACA16493	ACA16493 Prokaryot
c 519	32	72.7	78268	11	ACN44342	Acn44342 Human gen	c 592	31	70.5	211	10	ADBS2319	Adbs2319 Norway ra
c 520	32	72.7	86804	12	ADQ97700	Adq97700 Mouse can	c 593	31	70.5	211	10	ADH56106	Adh56106 Rat pain-
c 521	32	72.7	110000	2	AAK91990_10	Continuation (11 o	c 594	31	70.5	220	12	ADL10401	Adl10401 Cat flea
c 522	32	72.7	110000	3	AAF22305_05	Continuation (6 of	c 595	31	70.5	254	12	ADG00322	Adg00322 Nicotiana
c 523	32	72.7	110000	6	ABK08336_04	Continuation (5 of	c 596	31	70.5	272	5	AA568765	AA568765 DNA encod
c 524	32	72.7	110000	6	ABA90521_18	Continuation (19 o	c 597	31	70.5	301	4	AA550172	AA550172 Staphyloc
c 525	32	72.7	110000	8	AD552223_0	Ad552223 Human chr	c 598	31	70.5	301	8	ACA17765	ACA17765 Prokaryot
c 526	32	72.7	110000	10	ADF77343_09	Continuation (10 o	c 599	31	70.5	301	8	ACA17466	ACA17466 Prokaryot
c 527	32	72.7	110000	10	ACF65383_1	Continuation (2 of	600	31	70.5	301	8	ACA17406	ACA17406 Prokaryot
c 528	32	72.7	110000	10	ACF67367_02	Continuation (3 of	c 601	31	70.5	310	2	AAV78647	AAV78647 Staphyloc
c 529	32	72.7	110000	10	ACF67367_25	Continuation (26 o	c 602	31	70.5	320	12	ADP94337	Adp94337 Cotton ex
c 530	32	72.7	110000	10	ACF67367_38	Continuation (39 o	c 603	31	70.5	323	2	AAZ24584	Aaz24584 Human lun
c 531	32	72.7	110000	10	ACF65388_09	Continuation (10 o	c 604	31	70.5	323	2	AAZ24584	Aaz24584 Human lun
c 532	32	72.7	110000	10	ACF65388_10	Continuation (11 o	c 605	31	70.5	323	6	ABQ92228	Abq92228 Human lun
c 533	32	72.7	110000	10	ACF65386_3	Continuation (4 of	c 606	31	70.5	323	6	ABQ92228	Abq92228 Human lun
c 534	32	72.7	110000	11	ACN43998_3	Continuation (4 of	c 607	31	70.5	323	9	ADA28643	Ada28643 Human lun
c 535	32	72.7	110000	12	ADJ75985_04	Continuation (5 of	c 608	31	70.5	323	10	ADBS3603	Adbs3603 Human lun
c 536	32	72.7	110000	12	ADN97989_04	Continuation (5 of	c 609	31	70.5	323	10	ADH36738	Adh36738 Human lun
c 537	32	72.7	110000	12	ADO50281_04	Continuation (5 of	c 610	31	70.5	323	12	ADM56541	Adm56541 Human lun
c 538	32	72.7	110000	12	ADQ59443_1	Continuation (2 of	c 611	31	70.5	323	12	ADN89585	Adn89585 Human lun
c 539	32	72.7	110000	12	ADQ59443_2	Continuation (3 of	c 612	31	70.5	323	14	ADU98233	Adu98233 Lung tumo
c 540	32	72.7	110000	12	ADQ97050_3	Continuation (4 of	c 613	31	70.5	323	14	ABE10041	Aeb10041 Cancer re
c 541	32	72.7	110000	12	ADQ97960_0	Adq97960 Human can	c 614	31	70.5	334	2	AAV78632	AAV78632 Staphyloc
c 542	32	72.7	110000	14	ADZ13754_1	Continuation (2 of	c 615	31	70.5	346	2	AAV88599	AAV88599 SST clone
c 543	32	72.7	110000	14	ADZ13754_2	Continuation (3 of	c 616	31	70.5	375	11	ADM21656	Adm21656 Rat hepat
c 544	32	72.7	110000	14	ABE42737_09	Continuation (10 o	c 617	31	70.5	380	2	AAV78467	AAV78467 Staphyloc
c 545	32	72.7	110000	14	ABE85185_04	Continuation (5 of	c 618	31	70.5	385	4	AA550047	AA550047 Staphyloc
c 546	32	72.7	110218	11	ACN44744	Acn44744 Mouse gen	c 619	31	70.5	385	8	ACA17296	ACA17296 Prokaryot
c 547	32	72.7	110300	13	ADS36499	Ad36499 Human aut	c 620	31	70.5	398	2	AAV78310	AAV78310 Staphyloc
c 548	32	72.7	126266	11	ACN44602	Acn44602 Human gen	c 621	31	70.5	400	2	AAV77860	AAV77860 Staphyloc
c 549	32	72.7	132762	10	ADH63063	Adh63063 Human fib	c 622	31	70.5	400	2	AAV77875	AAV77875 Staphyloc
c 550	32	72.7	140342	14	ADZ13043	Adz13043 Human can	c 623	31	70.5	407	9	ACH47421	Ach47421 Human inf
c 551	32	72.7	148567	9	ACA62841	Ac62841 Human kin	c 624	31	70.5	409	14	ABE56232	Abe56232 HPV 67 DN
c 552	32	72.7	148567	10	ABSS5500	Abss5500 Gene enco	c 625	31	70.5	410	8	ABX62489	Abx62489 Arabidops
c 553	32	72.7	148567	12	ADL09163	Adl09163 Human pro	c 626	31	70.5	414	8	ABX62394	Abx62394 Arabidops
c 554	32	72.7	148567	14	ADZ58499	Adz58499 Human ser	c 627	31	70.5	434	9	ACH45995	Ach45995 Human inf
c 555	32	72.7	160755	4	AAH88704	Aah88704 Human DNA	c 628	31	70.5	442	6	ABL181756	Ab181756 Human ova
c 556	32	72.7	217409	11	ACN45150	Acn45150 Human gen	c 629	31	70.5	447	5	ADL73077	Adl73077 Human ova
c 557	32	72.7	252907	13	ABD32694	Abd32694 Human can	c 630	31	70.5	447	5	ADL38211	Adl38211 Human ova
c 558	32	72.7	254366	8	ABZ23704	Abz23704 Human pho	c 631	31	70.5	450	5	ABV10684	Abv10684 Human pro
c 559	32	72.7	263853	14	ABE39171	Ab39171 L. pneumo	c 632	31	70.5	461	2	AAV44706	AAV44706 Human pap
c 560	32	72.7	335199	10	ADC24703	Adc24703 Human wil	c 633	31	70.5	461	2	AAV78142	AAV78142 DNA seque
c 561	32	72.7	349980	5	AAH68531	Aah68531 C glutami	c 634	31	70.5	463	8	ACA15332	ACA15332 Prokaryot
c 562	32	72.7	349980	5	AAH68527	Aah68527 C glutami	c 635	31	70.5	463	8	ACA13740	ACA13740 Prokaryot
c 563	32	72.7	349980	5	AAH68532	Aah68532 C glutami	c 636	31	70.5	478	2	AAV78016	AAV78016 Staphyloc
c 564	32	72.7	349980	5	AAH41223	Aah41223 Pyrococcu	c 637	31	70.5	533	8	ABZ53477	Abz53477 Aspergill
c 565	31.5	71.6	2733	13	ADT46758	Adt46758 Bacterial	c 638	31	70.5	540	11	ABD11622	Abd11622 Pseudomon
c 566	31	70.5	50	6	ABZ00432	Abz00432 Human leu	c 639	31	70.5	540	11	ABD11622	Abd11622 Pseudomon
c 567	31	70.5	53	2	AAV79357	AAV79357 Staphyloc	c 640	31	70.5	543	3	AAV76812	AAV76812 Human ORF
c 568	31	70.5	156	4	AA550769	AA550769 Staphyloc	c 641	31	70.5	546	5	AAV53255	AAV53255 Human pro
c 569	31	70.5	156	8	ACA18052	ACA18052 Prokaryot	c 642	31	70.5	555	5	ADL44595	Adl44595 Human ova
c 570	31	70.5	158	12	ACH83327	Ach83327 Human gen	c 643	31	70.5	555	12	ACH69627	Ach69627 Human gen
c 571	31	70.5	162	8	ACA17554	ACA17554 Prokaryot	c 644	31	70.5	557	14	AAV90645	AAV90645 Human sec
c 572	31	70.5	162	8	ACA16376	ACA16376 Prokaryot	c 645	31	70.5	557	14	ACL63355	ACL63355 Human col
c 573	31	70.5	179	8	ACA17893	ACA17893 Prokaryot	c 646	31	70.5	581	2	AAV19018	AAV19018 Human gen
c 574	31	70.5	184	8	ACA17874	ACA17874 Prokaryot	c 647	31	70.5	587	13	ACN46126	Acn46126 Cotton pr
c 575	31	70.5	186	8	ACA17753	ACA17753 Prokaryot	c 648	31	70.5	600	6	ABN93226	Abn93226 Staphyloc
c 576	31	70.5	190	10	ADBE83290	Adbe83290 Bovine re	c 649	31	70.5	600	13	ADS04181	Ads04181 Staphyloc
c 577	31	70.5	190	12	ADI35805	Adi35805 Cow proop	c 650	31	70.5	605	4	AAV17157	AAV17157 Human ion
c 578	31	70.5	193	4	AA549964	AA549964 Staphyloc	c 651	31	70.5	605	8	ACD01542	ACD01542 cDNA clon
c 579	31	70.5	193	4	AA549985	AA549985 Staphyloc	c 652	31	70.5	605	10	ADBE29233	Adbe29233 Novel hum
c 580	31	70.5	193	4	AA550822	AA550822 Staphyloc	c 653	31	70.5	605	13	ACN46085	Acn46085 Cotton pr
c 581	31	70.5	193	4	AA549972	AA549972 Staphyloc	c 654	31	70.5	607	2	AAV36876	AAV36876 Piscicoli
c 582	31	70.5	193	8	ACA17208	ACA17208 Prokaryot	c 655	31	70.5	625	4	AAK88936	Aak88936 Human dig
c 583	31	70.5	193	8	ACA17207	ACA17207 Prokaryot	c 656	31	70.5	633	10	ADK54918	Adk54918 Plant DNA
c 584	31	70.5	193	8	ACA18096	ACA18096 Prokaryot	c 657	31	70.5	633	13	ADK63255	Adk63255 Cotton cd
c 585	31	70.5	193	8	ACA17210	ACA17210 Prokaryot	c 658	31	70.5	643	4	AA550027	AA550027 Staphyloc
c 586	31	70.5	201	4	AA550265	AA550265 Staphyloc	c 659	31	70.5	643	8	ACA17264	ACA17264 Prokaryot
c 587	31	70.5	201	8	ACA17522	ACA17522 Prokaryot	c 660	31	70.5	660	11	ACN82701	Acn82701 Breatst ca
c 588	31	70.5	201	13	ADS41219	Ads41219 Human aut	c 661	31	70.5	661	5	AAH87674	Aah87674 Peppermin

662	31	70.5	666	4	ABLI19029	Abli19029 Drosophil	c 735	31	70.5	1213	3	AAC49678	Aac49678 Arabidops
663	31	70.5	672	4	AAF71794	Aaf71794 Corynebac	c 736	31	70.5	1218	3	AAC42075	Aac42075 Arabidops
664	31	70.5	675	11	ACL27201	Acl27201 Rice abio	737	31	70.5	1224	12	ADJ53340	Adj53340 Amycolato
665	31	70.5	689	6	ABT08952	Abt08952 Phase-1 R	738	31	70.5	1238	13	ADM67823	Adm67823 Human sec
666	31	70.5	689	12	ADH22730	Adh22730 Partial D	739	31	70.5	1243	3	AAC40472	Aac40472 Arabidops
667	31	70.5	689	13	ADR91089	Adr91089 Spleen ne	740	31	70.5	1260	13	ADT14733	Adt14733 Plant CDN
668	31	70.5	693	5	ABA18176	Abal18176 Human ner	c 741	31	70.5	1293	8	ACA49289	Aca49289 Prokaryot
669	31	70.5	698	4	AAH08452	Aah08452 Human CDN	742	31	70.5	1295	6	ABK24530	Abk24530 BIF-2alph
670	31	70.5	699	3	AAF11905	Aaf11905 Aspergill	c 743	31	70.5	1296	11	ADJ11655	Adj11655 Rice DNA
671	31	70.5	699	13	ADU55946	Adu55946 Aspergill	744	31	70.5	1302	8	ACA23121	ACA23121 Prokaryot
672	31	70.5	699	14	ADZ93949	Adz93949 Aspergill	745	31	70.5	1302	13	ADU96768	Adu96768 Borrelia
673	31	70.5	703	6	ABT09324	Abt09324 Phase-1 R	c 746	31	70.5	1311	2	AAV77852	Aav77852 Staphyloc
674	31	70.5	703	10	ADG30831	Adg30831 Liver tox	c 747	31	70.5	1329	8	ACA18804	Aca18804 Prokaryot
675	31	70.5	731	4	AAI95610	Aai95610 Human neu	c 748	31	70.5	1350	13	ADT19545	Adt19545 Plant CDN
676	31	70.5	756	13	ADS50060	Ads50060 Bacterial	749	31	70.5	1355	2	AAK37390	Aak37390 Human sec
677	31	70.5	756	13	ADS55575	Ads55575 Bacterial	c 750	31	70.5	1356	10	ADE59548	Ade59548 Human gen
678	31	70.5	768	4	AAI95156	Aai95156 Human neu	c 751	31	70.5	1368	8	ACA51301	ACA51301 Prokaryot
679	31	70.5	777	13	ADS59978	Ads59978 Bacterial	c 752	31	70.5	1396	4	AAH55089	Aah55089 S. epider
680	31	70.5	792	4	AA822599	Aa822599 Human CDN	753	31	70.5	1416	13	ADS48225	Ads48225 Bacterial
681	31	70.5	792	4	AA822835	Aa822835 Human CDN	754	31	70.5	1431	14	ABE65088	Aeb65088 Rice geno
682	31	70.5	805	11	ACN85050	Acn85050 Breast ca	c 755	31	70.5	1442	6	ABS71606	Abs71606 Staphyloc
683	31	70.5	812	3	AAA87695	Aaa87695 Human sec	c 756	31	70.5	1468	6	ABS71609	Abs71609 Staphyloc
684	31	70.5	812	3	AAAC99855	Aac99855 Human sec	c 757	31	70.5	1475	6	ABS71607	Abs71607 Staphyloc
685	31	70.5	825	8	ACA25410	Aca25410 Prokaryot	c 758	31	70.5	1476	9	ABX11209	Abx11209 16S rRNA
686	31	70.5	858	2	AAZ15793	Aaz15793 Human gen	c 759	31	70.5	1476	10	ADC23320	Adc23320 DNA of 16
687	31	70.5	861	11	ACN82265	Acn82265 Breast ca	c 760	31	70.5	1484	5	AAE11026	Aae11026 Staphyloc
688	31	70.5	871	11	ACN84319	Acn84319 Breast ca	c 761	31	70.5	1490	2	AAK97626	Aak97626 Extended
689	31	70.5	880	11	ACN83185	Acn83185 Breast ca	c 762	31	70.5	1490	6	ABN87649	Abn87649 Staphyloc
690	31	70.5	887	4	AAJ49276	Aaj49276 Staphyloc	c 763	31	70.5	1490	12	ADP18893	Adp18893 Human sec
691	31	70.5	887	8	ACA18119	Aca18119 Prokaryot	c 764	31	70.5	1490	13	ADP55604	Adp55604 Human PRO
692	31	70.5	887	8	ACA16469	Aca16469 Prokaryot	c 765	31	70.5	1512	11	ABD02716	Abd02716 Pseudomon
693	31	70.5	893	8	AAJ51720	Aaj51720 Yeast int	c 766	31	70.5	1518	5	AAH65173	Aah65173 C glutami
694	31	70.5	893	9	ACC59725	Acc59725 Nucleotid	c 767	31	70.5	1519	14	ADV97729	Adv97729 CDNA sequ
695	31	70.5	906	11	ABN79846	Abn79846 Breast ca	c 768	31	70.5	1542	12	ADOS8171	Ados8171 S. epider
696	31	70.5	924	6	ABN83321	Abn83321 Human ren	c 769	31	70.5	1545	9	ABX11210	Abx11210 16S rRNA
697	31	70.5	950	8	ABX92256	Abx92256 Human ova	c 770	31	70.5	1545	10	ADC23321	Adc23321 DNA of 16
698	31	70.5	960	2	AAK91556	Aak91556 DNA encod	c 771	31	70.5	1551	9	ABX11208	Abx11208 16S rRNA
699	31	70.5	960	3	AAA81448	Aaa81448 N. mening	c 772	31	70.5	1551	10	ADC23319	Adc23319 DNA of 16
700	31	70.5	964	10	ADD45476	Add45476 Human gen	c 773	31	70.5	1554	9	ABX11211	Abx11211 16S rRNA
701	31	70.5	964	10	ADRS57445	Adrs57445 Human gen	c 774	31	70.5	1554	10	ADC23322	Adc23322 DNA of 16
702	31	70.5	972	6	ABL91249	Ab191249 Chlamydia	c 775	31	70.5	1555	2	AAAT29142	Aat29142 rRNA gene
703	31	70.5	972	13	ADR99384	Adr99384 Chlamydia	c 776	31	70.5	1555	2	AAV24294	Aav24294 Staphyloc
704	31	70.5	982	8	AAJ51710	Aaj51710 Yeast int	c 777	31	70.5	1555	6	AAK99056	Aak99056 DNA of th
705	31	70.5	982	9	ACC59702	Acc59702 Artificia	c 778	31	70.5	1555	6	AAK99099	Aak99099 DNA encod
706	31	70.5	985	13	ADT18840	Adt18840 Plant CDN	c 779	31	70.5	1555	9	ADB16301	Adb16301 Cleavage
707	31	70.5	1001	3	AAAC57441	Aac57441 Arachidon	c 780	31	70.5	1555	10	ADB61676	Adb61676 16S rRNA
708	31	70.5	1001	3	AAAC57440	Aac57440 Arachidon	c 781	31	70.5	1555	10	ADC02549	Adc02549 S. aureus
709	31	70.5	1001	3	AAAC57439	Aac57439 Arachidon	c 782	31	70.5	1555	14	ADW94287	Adw94287 Prolifera
710	31	70.5	1001	3	AAAC57438	Aac57438 Arachidon	c 783	31	70.5	1555	14	ADW94447	Adw94447 Prolifera
711	31	70.5	1001	3	AAAC57443	Aac57443 Arachidon	c 784	31	70.5	1555	14	ADW94575	Adw94575 Prolifera
712	31	70.5	1008	11	ABD11838	Abd11838 Pseudomon	c 785	31	70.5	1555	14	ADW94604	Adw94604 Prolifera
713	31	70.5	1021	3	AAAC40665	Aac40665 Arabidops	c 786	31	70.5	1555	14	ADW94509	Adw94509 Prolifera
714	31	70.5	1045	3	AAAC44329	Aac44329 Arabidops	c 787	31	70.5	1567	9	ABX11204	Abx11204 16S rRNA
715	31	70.5	1050	2	AAV49561	Aav49561 Human epi	c 788	31	70.5	1567	10	ADC23315	Adc23315 DNA of 16
716	31	70.5	1052	10	ADG37296	Adg37296 Aspergill	c 789	31	70.5	1574	12	ADM87016	Adm87016 Human pro
717	31	70.5	1066	13	ADRG3165	Adrg3165 Cottoon CD	c 790	31	70.5	1614	4	AAF61054	Aaf61054 P. putida
718	31	70.5	1078	2	AAZ96339	Aaz96339 S. pneumo	c 791	31	70.5	1641	9	ABX11202	Abx11202 16S rRNA
719	31	70.5	1088	14	ADY65705	Ady65705 S. mansoni	c 792	31	70.5	1641	10	ADC23313	Adc23313 DNA of 16
720	31	70.5	1093	11	ACN84270	Acn84270 Breast ca	c 793	31	70.5	1718	13	ADX13021	Adx13021 Plant ful
721	31	70.5	1130	14	ADW16720	Adw16720 Pinus rad	c 794	31	70.5	1728	8	ACA53857	Aca53857 Prokaryot
722	31	70.5	1146	12	ADJ34931	Adj34931 DNA encod	c 795	31	70.5	1734	4	AAK85087	Aak85087 Atheroscl
723	31	70.5	1150	14	ABE26911	Ab26911 Pinus rad	c 796	31	70.5	1734	11	ADM29575	Adm29575 Human ath
724	31	70.5	1161	9	ADA30101	Ada30101 DNA encod	c 797	31	70.5	1739	13	ADX65064	Adx65064 Plant ful
725	31	70.5	1162	3	AAAC34538	Aac34538 Arabidops	c 798	31	70.5	1785	2	AAV77317	Aav77317 Staphyloc
726	31	70.5	1170	8	AAAS52324	Aas52324 E. coli D	c 799	31	70.5	1788	2	AAK99158	Aak99158 DNA encod
727	31	70.5	1170	8	AAAC32385	Aac32385 Prokaryot	c 800	31	70.5	1793	12	ADL22886	Adl22886 Human MP2
728	31	70.5	1176	13	ADG45979	Adg45979 Bacterial	c 801	31	70.5	1801	12	ADG18921	Adg18921 Human KRI
729	31	70.5	1183	3	ADF00794	Adf00794 Bacterial	c 802	31	70.5	1801	12	ADN05399	Adn05399 Antipsori
730	31	70.5	1183	3	AAAC1302	Aac1302 Arabidops	c 803	31	70.5	1805	8	ACC69468	Acc69468 Human mal
731	31	70.5	1191	8	ACA27583	Aca27583 Prokaryot	c 804	31	70.5	1805	8	ACC69481	Acc69481 Human mal
732	31	70.5	1197	13	ADSA7832	Ads7832 Bacterial	c 805	31	70.5	1812	4	AAAD08314	Aad08314 Human sec
733	31	70.5	1206	4	AAAS51687	Aas51687 Staphyloc	c 806	31	70.5	1834	3	AAAD00681	Aad00681 Human Hyd
734	31	70.5	1207	8	ACA21713	Aca21713 Prokaryot	c 807	31	70.5	1860	2	AAK99157	Aak99157 DNA encod

808	31	70.5	1862	2	AAZ224474	Aaz24474 Spinach g	C 881	31	70.5	2538	8	ACA50540	Aca50540 Prokaryot
809	31	70.5	1885	13	ADX61150	Adx61150 Plant ful	C 882	31	70.5	2595	6	ABN67080	Abn67080 Streptoco
810	31	70.5	1894	2	AAQ61796	Aaq61796 Human NRS	C 883	31	70.5	2598	4	AAH17483	Aah17483 Human CDN
811	31	70.5	1894	2	AAQ61796	Aaq61796 Human NRS	C 884	31	70.5	2598	4	AAH17483	Aah17483 Human CDN
812	31	70.5	1902	5	AAQ89042	Aaq89042 Rat NDF c	C 885	31	70.5	2598	5	AAI93859	Aai93859 Human sto
813	31	70.5	1905	8	AAQ89042	Aaq89042 Rat NDF c	C 886	31	70.5	2598	12	ADQ85658	Adq85658 Human tum
814	31	70.5	1907	8	AAV84508	Aav84508 Human Plk	C 887	31	70.5	2603	11	ADM03641	Adm03641 Human CDN
815	31	70.5	1907	4	AAV84508	Aav84508 Human Plk	C 888	31	70.5	2666	4	ABL19028	AbL19028 Drosophil
816	31	70.5	1907	4	AAV84508	Aav84508 Human Plk	C 889	31	70.5	2683	4	ABL19030	AbL19030 Drosophil
817	31	70.5	1907	8	AAV84508	Aav84508 Human Plk	C 890	31	70.5	2742	4	ABL19030	AbL19030 Drosophil
818	31	70.5	1907	8	AAV84508	Aav84508 Human Plk	C 891	31	70.5	2750	4	ABL19030	AbL19030 Drosophil
819	31	70.5	1907	9	ACD44602	AcD44602 Human sec	C 892	31	70.5	2783	4	AAH55014	Aah55014 S. epider
820	31	70.5	1907	9	ACD44602	AcD44602 Human sec	C 893	31	70.5	2820	4	AAH55014	Aah55014 S. epider
821	31	70.5	1907	9	ACD44602	AcD44602 Human sec	C 894	31	70.5	2821	4	ABL08992	AbL08992 Drosophil
822	31	70.5	1917	5	ADL63571	AdL63571 Human ova	C 895	31	70.5	2839	4	AAH54998	Aah54998 S. epider
823	31	70.5	1923	4	ABL28551	AbL28551 Drosophil	C 896	31	70.5	2902	4	AAH18271	Aah18271 Human CDN
824	31	70.5	1937	2	AAZ00460	Aaz00460 Human sec	C 897	31	70.5	2904	13	ADX34642	Adx34642 Plant ful
825	31	70.5	1937	3	AAZ00460	Aaz00460 Human sec	C 898	31	70.5	2914	8	AAH51403	Aah51403 Human mic
826	31	70.5	1944	4	ABL02947	AbL02947 Human sec	C 899	31	70.5	2914	10	ADG91741	Adg91741 Human mic
827	31	70.5	1947	12	ADJ39852	AdJ39852 Plant CDN	C 900	31	70.5	2914	12	ADI29091	Adi29091 Human CDN
828	31	70.5	1951	4	AAZ00882	Aaz00882 Human sec	C 901	31	70.5	2967	4	AAH54254	Aah54254 S. epider
829	31	70.5	1952	4	AAZ00882	Aaz00882 Human sec	C 902	31	70.5	2998	14	ADY20432	Ady20432 DNA encod
830	31	70.5	1963	3	AAZ00882	Aaz00882 Human sec	C 903	31	70.5	3024	5	AAH50068	Aah50068 DNA encod
831	31	70.5	1977	3	AAZ00882	Aaz00882 Human sec	C 904	31	70.5	3024	5	AAH50068	Aah50068 DNA encod
832	31	70.5	2010	2	AAZ40848	Aaz40848 Arabidops	C 905	31	70.5	3088	7	ADY36408	Ady36408 HIRA geno
833	31	70.5	2010	11	ADM77870	Adm77870 Human CDN	C 906	31	70.5	3088	7	ADY36408	Ady36408 HIRA geno
834	31	70.5	2010	12	ADP19146	AdP19146 Human sec	C 907	31	70.5	3117	5	AAH50916	Aah50916 DNA encod
835	31	70.5	2010	14	ADZ89330	Adz89330 Secreted	C 908	31	70.5	3133	14	ABE71304	Aeb71304 Sequence
836	31	70.5	2019	8	ACC69469	Acc69469 Human mal	C 909	31	70.5	3169	4	AAH53992	Aah53992 S. epider
837	31	70.5	2033	2	AAV49560	Aav49560 Human epi	C 910	31	70.5	3171	4	AAH54173	Aah54173 S. epider
838	31	70.5	2034	6	AAV49560	Aav49560 Human epi	C 911	31	70.5	3191	5	ABV23079	Abv23079 Human pro
839	31	70.5	2044	13	ADZ67118	AdZ67118 Human bla	C 912	31	70.5	3191	5	ABV23079	Abv23079 Human pro
840	31	70.5	2044	14	ADZ67118	AdZ67118 Human bla	C 913	31	70.5	3200	11	ADM02274	Adm02274 Human CDN
841	31	70.5	2044	14	ADZ67118	AdZ67118 Human bla	C 914	31	70.5	3204	10	ADE56923	AdE56923 Rat gene
842	31	70.5	2044	14	ABE56457	Abe56457 Radiochem	C 915	31	70.5	3204	13	ADY07656	Ady07656 Pull leng
843	31	70.5	2044	14	ABE56457	Abe56457 Radiochem	C 916	31	70.5	3253	4	AAH54591	Aah54591 S. epider
844	31	70.5	2049	7	ADZ73147	AdZ73147 Human kid	C 917	31	70.5	3275	4	AAH54193	Aah54193 S. epider
845	31	70.5	2049	7	ADZ73147	AdZ73147 Human kid	C 918	31	70.5	3294	2	AAH54193	Aah54193 S. epider
846	31	70.5	2049	10	ADG32876	AdG32876 Human DNA	C 919	31	70.5	3294	2	AAH54193	Aah54193 S. epider
847	31	70.5	2078	2	AAV84627	Aav84627 Human sec	C 920	31	70.5	3300	3	AAH54541	Aah54541 S. epider
848	31	70.5	2084	4	ABA83417	AbA83417 Human sec	C 921	31	70.5	3308	4	AAH54779	Aah54779 S. epider
849	31	70.5	2084	8	ACC50830	Acc50830 Human sec	C 922	31	70.5	3310	4	AAH54779	Aah54779 S. epider
850	31	70.5	2084	8	ABZ71466	AbZ71466 Secreted	C 923	31	70.5	3337	12	ADQ63058	Adq63058 Novel hum
851	31	70.5	2084	9	ACH04918	Ach04918 Novel hum	C 924	31	70.5	3427	4	AAH54520	Aah54520 S. epider
852	31	70.5	2084	9	ACD44728	AcD44728 Human sec	C 925	31	70.5	3444	4	AAH54992	Aah54992 S. epider
853	31	70.5	2084	9	ADB91427	AdB91427 Human sec	C 926	31	70.5	3473	13	ADR66617	Adr66617 Human pro
854	31	70.5	2084	10	ADC73942	AdC73942 Human sec	C 927	31	70.5	3473	13	ADR66617	Adr66617 Human pro
855	31	70.5	2092	5	AAZ93774	Aaz93774 Human CDN	C 928	31	70.5	3500	4	AAH54518	Aah54518 S. epider
856	31	70.5	2092	14	ADY63058	Ady63058 Human cto	C 929	31	70.5	3513	4	AAH53987	Aah53987 S. epider
857	31	70.5	2099	4	AAZ93774	Aaz93774 Human CDN	C 930	31	70.5	3540	12	ADI82461	Adi82461 Human mod
858	31	70.5	2108	4	AAZ93774	Aaz93774 Human CDN	C 931	31	70.5	3540	14	ADY79900	Ady79900 Nucleotid
859	31	70.5	2109	10	ADC37350	AdC37350 Nuclear f	C 932	31	70.5	3581	10	ADB58164	AdB58164 Toxicity-
860	31	70.5	2119	6	ABL89563	AbL89563 Human pol	C 933	31	70.5	3581	10	ADB58164	AdB58164 Toxicity-
861	31	70.5	2130	11	ABD02544	AbD02544 Pseudomon	C 934	31	70.5	3586	4	AAH16064	Aah16064 Human CDN
862	31	70.5	2133	2	AAZ44121	Aaz44121 Wild type	C 935	31	70.5	3586	5	AAH16064	Aah16064 Human CDN
863	31	70.5	2133	3	AAZ44121	Aaz44121 Wild type	C 936	31	70.5	3625	10	AAH16064	Aah16064 Human CDN
864	31	70.5	2137	10	ADZ68430	AdZ68430 Lolium pe	C 937	31	70.5	3657	4	AAH54823	Aah54823 S. epider
865	31	70.5	2137	14	ABE03133	AbE03133 Fructan b	C 938	31	70.5	3713	4	AAH54014	Aah54014 S. epider
866	31	70.5	2138	10	ADZ68533	AdZ68533 Lolium pe	C 939	31	70.5	3724	4	ABL02952	AbL02952 Drosophil
867	31	70.5	2138	14	ABE03018	AbE03018 Fructan b	C 940	31	70.5	3821	4	AAH54603	Aah54603 S. epider
868	31	70.5	2186	2	AAQ51344	Aaq51344 Rat NRSP	C 941	31	70.5	3842	10	ADK67002	Adk67002 Gene #92
869	31	70.5	2186	2	AAQ51344	Aaq51344 Rat NRSP	C 942	31	70.5	3842	14	ADY98482	Ady98482 Human mye
870	31	70.5	2186	4	ABA08649	AbA08649 Human sec	C 943	31	70.5	3878	10	ADD01253	AdD01253 Human nuc
871	31	70.5	2205	13	AAZ56977	Aaz56977 Arabidops	C 944	31	70.5	3929	5	AAH54246	Aah54246 S. epider
872	31	70.5	2245	3	ADZ64420	AdZ64420 Plant ful	C 945	31	70.5	3990	5	ABL61803	AbL61803 Human ova
873	31	70.5	2281	11	ADI30700	Adi30700 Human CDN	C 946	31	70.5	4066	4	ABL02946	AbL02946 Drosophil
874	31	70.5	2281	13	ADZ82768	AdZ82768 Human lym	C 947	31	70.5	4324	4	ABL28550	AbL28550 Drosophil
875	31	70.5	2291	2	AAZ99133	Aaz99133 DNA encod	C 948	31	70.5	4429	4	AAH54300	Aah54300 S. epider
876	31	70.5	2297	12	ADQ23781	AdQ23781 Human sof	C 949	31	70.5	4442	12	ADQ63021	Adq63021 Novel hum
877	31	70.5	2329	8	ABZ63416	AbZ63416 Human CDN	C 950	31	70.5	4707	12	ADQ64558	Adq64558 Novel hum
878	31	70.5	2350	8	ACC69465	Acc69465 Human mal	C 951	31	70.5	4869	4	ABL21338	AbL21338 Drosophil
879	31	70.5	2393	6	ABQ99285	AbQ99285 Human cod	C 952	31	70.5	4996	14	ADY16350	Ady16350 DNA encod
880	31	70.5	2474	6	ABQ73247	AbQ73247 Lolium pe	C 953	31	70.5	4996	14	ABE81215	Abe81215 Human lym

```

c 954 31 70.5 5029 14 ADW94255
955 31 70.5 5029 14 ADY16352
956 31 70.5 5029 14 ADY15937
957 31 70.5 5029 14 AEA81214
c 958 31 70.5 5134 14 ADW94228
c 959 31 70.5 5648 14 ADW94202
960 31 70.5 5809 10 ADA19384
c 961 31 70.5 5904 4 ADL18535
962 31 70.5 5948 4 AAK87044
963 31 70.5 5948 4 AAK68700
c 964 31 70.5 5973 14 ADW94079
c 965 31 70.5 6048 8 ABQ78298
966 31 70.5 6048 8 ACA03916
c 967 31 70.5 6048 8 ACA41718
c 968 31 70.5 6048 14 AEB92043
c 969 31 70.5 6242 4 AAS27776
c 970 31 70.5 6242 10 ADB94579
c 971 31 70.5 7071 6 ABL61772
c 972 31 70.5 7071 6 ABL66571
c 973 31 70.5 7071 14 ADX08008
974 31 70.5 7253 3 AA92499
c 975 31 70.5 7291 14 ADW94182
976 31 70.5 7326 4 ABL10829
977 31 70.5 7326 13 ADQ99639
978 31 70.5 7572 14 AEB1817
979 31 70.5 7769 4 ABL30098
c 980 31 70.5 8247 4 ABL03218
981 31 70.5 8317 10 AAS35779
982 31 70.5 8317 4 ADE46473
983 31 70.5 8317 13 ADJ07891
984 31 70.5 8685 4 ABL18534
c 985 31 70.5 8734 2 AAX27723
986 31 70.5 10072 4 ABL10828
987 31 70.5 10242 4 ABL29075
988 31 70.5 11425 4 AAK79581
c 989 31 70.5 11881 4 AAS27693
c 990 31 70.5 11881 4 AAS27691
c 991 31 70.5 11881 4 AAS36624
c 992 31 70.5 11881 10 ADB94494
c 993 31 70.5 11881 10 ADB94496
c 994 31 70.5 11881 10 ADB94496
c 995 31 70.5 11881 10 ADE47318
c 996 31 70.5 11881 13 ADJ08736
c 997 31 70.5 12685 3 AA50043
c 998 31 70.5 12685 12 ADN07622
c 999 31 70.5 12685 14 ADW79939
1000 31 70.5 13382 14 ADW44488

```

ALIGNMENTS

```

RESULT 1
ABX40605
ID ABX40605 standard; cDNA; 343 BP.
XX
AC ABX40605;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #5770.
XX
KW Bovine; es; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
FN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PP 24-SEP-2001; 2001US-00960352.
XX

```

```

PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
DR
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 5770; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived from
XX cattle, and the LMFD nucleic acid can specifically hybridise to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
XX acid linked to a promoter and a 3' non- translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMFD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 343 BP; 40 A; 146 C; 108 G; 49 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.23 Length: 343
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x ABX40605 (1-343)
QY 1 PheLeuThrGlyAaGlnLeuAlaVal 9
Db 179 TTCTCAGCGGCAACCCAGCTGCGCGTG 205

RESULT 2
ACLS56146
ID ACL56146 standard; cDNA; 505 BP.
XX
XX ACL56146;
XX
XX 24-MAR-2005 (first entry)
DT
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:2281.
DE
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;

```

KW ss.
 XX Homo sapiens.
 OS WO200500087-A2.
 FN 06-JAN-2005.
 XX 13-MAY-2004; 2004WO-US015421.
 XX 03-JUN-2003; 2003US-0475872P.
 XX (CHIR) CHIRON CORP.
 XX Randazzo F, Moler B, Escobedo J, Garcia PD;
 XX WPI; 2005-075421/08.
 XX New isolated polynucleotides, which are differentially expressed in colon
 XX cancer cell, useful for treating cancer, e.g. colon cancer, breast
 XX cancer, or pancreatic cancer.
 XX Claim 1; SEQ ID NO 2281; 97pp; English.
 XX The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
 XX are differentially expressed in colon cancer cells. The invention also
 XX relates to vectors and host cells comprising a differentially expressed
 XX polynucleotide of the invention; a method for detecting a cancerous cell
 XX by detection of a gene product of the polynucleotides; a method for
 XX inhibiting a cancerous phenotype of a cell by inhibiting a gene product
 XX of the polynucleotides; a method of treating an individual with cancer by
 XX administration of a modulator of a gene product of the polynucleotides;
 XX and an isolated antibody that specifically binds to a polypeptide encoded
 XX by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
 XX antibodies, and methods are useful for the detection of cancerous cells;
 XX for the diagnosis, prognosis and management of cancer; for the
 XX identification of agents that modulate the phenotype of cancerous cells;
 XX for the identification of therapeutic targets for cancer chemotherapy;
 XX and for the treatment of cancer, especially colon cancer and metastasized
 XX colon cancer, but also breast or pancreatic cancer. The polynucleotides
 XX are also useful as a source of probes or primers for use in diagnostic
 XX methods. The differentially expressed polynucleotides or their encoded
 XX proteins can additionally be used as vaccines to modulate primary immune
 XX responses for the prevention or treatment of cancer. The present sequence
 XX represents a specifically claimed polynucleotide which is differentially
 XX expressed in colon cancer. Note: The sequence data for this patent did
 XX not form part of the printed specification, but was obtained in
 XX electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 505 BP; 68 A; 202 C; 149 G; 85 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 1.93 Length: 505
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0
 US-10-774-176-5 (1-9) x ACL56146 (1-505)
 Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 180 TTCCTACCGGACACCTGGCCGTG 206
 RESULT 3
 ABK87175
 ID ABK87175 standard; cDNA; 1260 BP.
 XX
 AC ABK87175;
 XX
 DT 07-OCT-2002 (first entry)

XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
 DE
 XX Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX
 OS Felis sp.
 XX
 XX Location/Qualifiers
 Key 1..1260
 CDS /*tag= a
 FT /product= "5T4 protein"
 FT
 XX WO200238612-A2.
 XX 16-MAY-2002.
 XX 13-NOV-2001; 2001WO-GB005004.
 XX 13-NOV-2000; 2000WO-GB004317.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Myers K, Drury N, Carroll M;
 XX WPI; 2002-557449/59.
 XX P-PSDB; AAU98694.
 XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 XX polypeptide, useful in preparation of vaccine for treating and/or
 XX preventing cancer in a subject, preferably a dog or cat.
 XX Claim 4; Page 68; 68pp; English.
 XX The present invention relates to the isolation of canine and feline
 XX oncofoetal leucine-rich glycoproteins known as 5T4, and the
 XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 XX a significant proportion of tumours. The sequences of the invention are
 XX useful in a pharmaceutical composition for the prevention and/or
 XX treatment of tumours or other diseases associated with cell
 XX proliferation, infections, and inflammatory conditions in animals,
 XX preferably dogs or cats. The compositions may also be used for cancer
 XX immunotherapy in these animals. The sequences of the invention may also
 XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
 XX measurement and localisation of 5T4 in extracts of plasma, urine,
 XX tissues, and in cell culture media. Antibodies specific for the 5T4
 XX protein are useful for isolating foetal cells from maternal blood. The
 XX isolation process may form part of a diagnostic method e.g. the foetal
 XX cells may then be subject to biochemical or genetic sampling used for
 XX testing foetal abnormalities, or to determine the sex of the foetus(ss).
 XX The present sequence encodes feline 5T4 protein
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 5.62 Length: 1260
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-5 (1-9) x ABK87175 (1-1260)
 Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 286 TTCCTACCGGCAATCAGCTGGCCGTG 312
 RESULT 4
 ID ABK87175
 DT ADB97513 standard; DNA; 1260 BP.

XX ADB97513;
AC
XX
DT 04-DEC-2003 (first entry)
XX
XX Feline 574 antigen DNA.
XX
XX Major Histocompatibility Complex class I peptide epitope; MHC;
KW 574 antigen; 574 epitope; polypeptide string; vaccine; T cell;
KW cytostatic; cancer; feline; gene; ds.
XX
XX Unidentified.
XX
XX
XX Key Location/Qualifiers
FT CDS 1..1260
FT /*tag= a
FT /product= "Feline 574 antigen protein"
XX
XX WO2003068816-A1.
XX
XX 21-AUG-2003.
XX
XX 13-FEB-2003; 2003WO-GB000670.
XX
XX 13-FEB-2002; 2002GB-00003419.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll M, Kingman S, Redchenko I;
XX
XX WPI, 2003-637141/60.
DR P-PSDB; ADB97520.
XX
XX New major histocompatibility complex class I peptide epitopes from human
PT 574 tumor-associated antigen, useful for preventing and/or treating a
PT disease, particularly cancer.
XX
XX Disclosure; Page 67; 73pp; English.
XX
XX The invention relates to a novel Major Histocompatibility Complex (MHC)
CC class I peptide epitope of the 574 antigen. The invention further
CC provides a polypeptide string comprising the 574 epitope; a nucleic acid
CC sequence encoding the 574 epitope or a polypeptide string of the 574
CC epitope; a vector system capable of delivering the 574 epitope nucleic
CC acid to a cell; a cell pulsed with the 574 epitope, a polypeptide of the
CC 574 epitope, its encoding nucleic acid, or the vector system; a vaccine
CC comprising the above; a method for treating and/or preventing a disease
CC in a subject by administering the vaccine; an agent capable of binding
CC specifically to the 574 epitope and/its encoding nucleic acid; a method
CC comprising detecting the presence of the 574 epitope or its encoding
CC nucleic acid in a subject; and a T cell line or clone capable of
CC specifically recognising the 574 epitope in conjunction with an MHC class
CC I molecule. The 574 epitope has cytostatic activity. The vaccine
CC comprising the 574 epitope or its encoding nucleic acid and the vector
CC system or cell is useful in the prevention and/or treatment of a disease,
CC particularly cancer. The detection method is useful for diagnosing or
CC monitoring the progression of a cancerous disease, and for detecting the
CC presence of the 574 epitope or its nucleic acid. The T cell line or clone
CC is useful in the manufacture of a medicament for treating and/or
CC preventing a disease. This polynucleotide sequence represents the feline
CC 574 antigen coding DNA of the invention.
XX
SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 5.62 Length: 1260
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-5 (1-9) x ADB97513 (1-1260)

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
|||||
Db 286 TTCTCACCAGCAGTCAGTCGCGTG 312
RESULT 5
ADB97452
ID ADB97452 standard; DNA; 1260 BP.
XX
XX ADB97452;
XX
XX 04-DEC-2003 (first entry)
XX
XX DNA encoding feline 574 protein.
XX
XX gene; ds; feline; Major Histocompatibility Complex class II; MHC;
KW epitope; 574 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FT CDS 1..1260
FT /*tag= a
FT /product= "Feline 574 antigen protein"
XX
XX WO2003068815-A2.
XX
XX 21-AUG-2003.
XX
XX 13-FEB-2003; 2003WO-GB000618.
XX
XX 13-FEB-2002; 2002GB-00003420.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll M, Harrop R, Kingsman S;
XX
XX WPI, 2003-663795/62.
DR P-PSDB; ADB97455.
XX
XX New Major Histocompatibility Complex class II peptide epitope of 574,
PT useful for manufacturing a medicament for diagnosing, preventing and/or
PT treating a disease, e.g. cancer.
XX
XX Disclosure; Page 49; 63pp; English.
XX
XX The invention relates to a Major Histocompatibility Complex (MHC) class
CC II peptide epitope of the 574 antigen. The vaccine or T-cell line or
CC clone has a cytostatic activity, as it is useful in manufacturing a
CC medicament for preventing and/or treating a disease, particularly cancer.
CC The methods are useful for detecting T-cells capable of specifically
CC recognising a peptide epitope in conjunction with an MHC molecule, for
CC diagnosing or monitoring the progression of a cancerous disease, or for
CC detecting the presence of a peptide or nucleic acid using an agent. The
CC MHC class II peptide epitope of the invention can be used in gene therapy
CC or as part of a vaccine. This polynucleotide sequence represents the DNA
CC coding for the feline 574 protein.
XX
SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 5.62 Length: 1260
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-5 (1-9) x ADB97452 (1-1260)
QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
|||||
Db 286 TTCTCACCAGCAGTCAGTCGCGTG 312

RESULT 6
 AAA27058
 ID AAA27058 standard; DNA; 1263 BP.
 XX
 AC AAA27058;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human 5T4 tumour-associated antigen gene.
 XX
 KW Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
 KW ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200029428-A2.
 XX
 PD 25-MAY-2000.
 XX
 PF 18-NOV-1999; 99WO-GB003859.
 XX
 PR 18-NOV-1998; 98GB-00025303.
 XX
 PR 27-JAN-1999; 99GB-00001739.
 XX
 PR 30-JUL-1999; 99GB-00017995.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Carroll MW, Myers KA;
 XX
 DR WPI; 2000-387735/33.
 XX
 PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.
 XX
 PS Example 2; Page 78; 79pp; English.
 XX
 CC The present sequence encodes the human 5T4 tumour-associated antigen
 CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
 CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The 5T4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject
 XX
 SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.63 Length: 1263
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-5 (1-9) x AAA27058 (1-1263)
 QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 DB 289 TTCTTACCGGCAACCGCTGGCGGTG 315
 RESULT 7
 AAF89736
 ID AAF89736 standard; DNA; 1263 BP.
 XX
 AC AAF89736;
 XX
 DT 23-JUL-2001 (first entry)

XX Nucleotide sequence of canine 5T4 protein.
 DE
 KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
 KW hypersensitivity; autoimmune disease; central nervous system disorder;
 KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
 KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
 KW Helicobacter-related disease; immune disorder; ss.
 XX
 OS Canis sp.
 XX
 FH Key Location/Qualifiers
 FT 1..1263
 FT /*tag= a
 FT /product= "5T4"
 XX
 PN WO200136486-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 13-NOV-2000; 2000WO-GB004317.
 XX
 PR 18-NOV-1999; 99WO-GB003859.
 XX
 PR 15-FEB-2000; 2000GB-00003527.
 XX
 PR 02-MAR-2000; 2000GB-00005071.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
 PI Myers KA;
 XX
 DR WPI; 2001-343805/36.
 XX
 DR P-PSDB; AAB83839.
 XX
 PT Use of single chain antibody capable of recognizing a disease associated
 PT molecule for manufacturing a medicament for preventing and/or treating a
 PT disease condition associated with disease associated molecule.
 XX
 PS Disclosure; Fig 26; 118pp; English.
 XX
 CC The specification describes the use of a single chain antibody (ScFv),
 CC which is capable of recognizing a disease associated molecule in the
 CC manufacture of a medicament for the prevention and treatment of a disease
 CC condition. The ScFv antibody is useful in the manufacture of a
 CC medicament, for affecting a disease in vivo, for preparing a
 CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
 CC treatment of a disease. The ScFv antibody is also useful for treating
 CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
 CC diseases, cancers, central nervous system disorders including Parkinson's
 CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
 CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
 CC related diseases, and other immune disorders. The present sequence
 CC encodes a 5T4 protein, which is used to produce ScFv of the invention
 XX
 SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.63 Length: 1263
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-5 (1-9) x AAF89736 (1-1263)
 QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 DB 289 TTCTTACCGGCAACCGCTGGCGGTG 315
 RESULT 8
 ABK87174
 ID ABK87174 standard; cDNA; 1263 BP.

```

XX AC ABK87174;
XX DT 07-OCT-2002 (first entry)
XX DB cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
XX KW Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
XX KW cell proliferative disorder; infection; inflammatory condition;
XX KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
XX KW foetal abnormality; foetal sex determination; gene; ss.
XX OS Canis sp.
XX PH Key Location/Qualifiers
XX FT CDS 1..1263
XX FT /*tag= a
XX FT /product= "5T4 protein"
XX PN WO200238612-A2.
XX XX 16-MAY-2002.
XX XX 13-NOV-2001; 2001WO-GB005004.
XX XX 13-NOV-2000; 2000WO-GB004317.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Myers K, Drury N, Carroll M;
XX DR WPI; 2002-557449/59.
XX DR P-PSDB; AAU98693.
XX PT Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
XX PT polypeptide, useful in preparation of vaccine for treating and/or
XX PT preventing cancer in a subject, preferably a dog or cat.
XX PS Claim 1; Page 67; 68pp; English.
XX CC The present invention relates to the isolation of canine and feline
XX CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
XX CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
XX CC a significant proportion of tumours. The sequences of the invention are
XX CC useful in a pharmaceutical composition for the prevention and/or
XX CC treatment of tumours or other diseases associated with cell
XX CC proliferation, infections, and inflammatory conditions in animals,
XX CC preferably dogs or cats. The compositions may also be used for cancer
XX CC immunotherapy in these animals. The sequences of the invention may also
XX CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
XX CC measurement and localisation of 5T4 in extracts of plasma, urine,
XX CC tissues, and in cell culture media. Antibodies specific for the 5T4
XX CC protein are useful for isolating foetal cells from maternal blood. The
XX CC isolation process may form part of a diagnostic method e.g. the foetal
XX CC cells may then be subject to biochemical or genetic sampling used for
XX CC testing foetal abnormalities, or to determine the sex of the foetus(es).
XX CC The present sequence encodes canine 5T4 protein
XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.63 Length: 1263
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x ABK87174 (1-1263)
QY 1 PheLeuThrGlyAaGlnLeuAlaVal 9
DB 289 TTCTTACGGGCAACCAAGCTGGCGTG 315

XX AC AAD56199 standard; DNA; 1331 BP.
XX DT 07-AUG-2003 (first entry)
XX DE Human LRRCAPS related DNA #6.
XX KW Human; p53 pathway; Leucine rich repeat capricious related protein;
XX KW LRRCAPS; cancer; gene therapy; ds.
XX OS Homo sapiens.
XX PN WO2003035831-A2.
XX PD 01-MAY-2003.
XX PF 21-OCT-2002; 2002WO-US033540.
XX PR 22-OCT-2001; 2001US-0338733P.
XX PR 15-FEB-2002; 2002US-0357600P.
XX PR 01-MAR-2002; 2002US-0361196P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
XX PI Francie-Lang H, Friedman L;
XX DR WPI; 2003-421410/39.
XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
XX PT comprises contacting an assay system comprising a purified leucine rich
XX PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX PS Disclosure; Page 75-76; 99pp; English.
XX CC The invention relates to a method of identifying a candidate p53 pathway
XX CC modulating agent. The method involves contacting an assay system
XX CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX CC polypeptide or nucleic acid or its fragment with a test agent and
XX CC detecting a test agent-biased activity, where a difference between the
XX CC test agent-biased activity and the reference activity identifies the test
XX CC agent as a candidate p53 pathway modulating agent. The method is useful
XX CC for identifying a candidate p53 pathway-modulating agent for preparing a
XX CC composition for diagnosing or treating cancer. The invention is useful in
XX CC gene therapy. The present sequence is human LRRCAPS related DNA
XX SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.99 Length: 1331
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x AAD56199 (1-1331)
QY 1 PheLeuThrGlyAaGlnLeuAlaVal 9
DB 319 TTCTTACGGGCAACCAAGCTGGCGTG 345

XX AC ADJ56299 standard; cDNA; 2020 BP.
XX DT 06-MAY-2004 (first entry)

```

XX DE Human cDNA differentially expressed in MYCN activated cells SeqID 105.
XX KW human; differential expression; transactivator; proto-oncogene;
KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;
KW MYCN activated cell.
XX OS Homo sapiens.
XX PA US2003119009-A1.
XX PN 26-JUN-2003.
XX PD 25-FEB-2002; 2002US-00084817.
XX PF 23-FEB-2001; 2001US-0270784P.
XX PR (STUA//) STUART S G.
XX PA (NUCH//) NUCHTERN J G.
XX PA (PLON//) PLON S E.
XX PA (SHOH//) SHOHET J M.
XX PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX PI WPI; 2003-635698/60.
XX DR New genes regulated by MYCN activation, useful in gene therapy,
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX PS Claim 1; SEQ ID NO 105; 27pp; English.
XX CC This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20030119009.
XX SQ Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 9.74 Length: 2020
XX Score: 44.00 Matches: 9
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 100.0% Indels: 0
XX DB: 10 Gaps: 0
XX
XX US-10-774-176-5 (1-9) x ADJ56299 (1-2020)
XX QY 1 PheLeuThrGlyAenGlnLeuAlaVal 9
XX Db 359 TTCTTTACCGGACACCGAGTGGCGTG 385
XX
XX RESULT 11
XX ACC51052
XX ID ACC51052 standard; cDNA; 2053 BP.
XX AC ACC51052;
XX XX

DT 12-JUN-2003 (first entry)
XX Human bladder cancer associated cDNA sequence SEQ ID NO:192.
XX KW Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.
XX OS Homo sapiens.
XX PN WO2003003906-A2.
XX PD 16-JAN-2003.
XX PF 03-JUL-2002; 2002WO-US021338.
XX PR 03-JUL-2001; 2001US-0302814P.
XX PR 03-AUG-2001; 2001US-0310099P.
XX PR 08-NOV-2001; 2001US-0343705P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Mack DH, Aziz N;
XX PI WPI; 2003-201532/19.
XX DR P-PSDB; ABR48236.
XX PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX PS Claim 6; Page 296; 307pp; English.
XX CC The present invention describes a method for detecting a bladder cancer-
XX associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridises to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 9.93 Length: 2053
XX Score: 44.00 Matches: 9
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 100.0% Indels: 0
XX DB: 8 Gaps: 0
XX
XX US-10-774-176-5 (1-9) x ACC51052 (1-2053)
XX QY 1 PheLeuThrGlyAenGlnLeuAlaVal 9
XX Db 373 TTCTTTACCGGACACCGAGTGGCGTG 399
XX
XX RESULT 12
XX ABX76332
XX ID ABX76332 standard; DNA; 2053 BP.
XX AC ABX76332;
XX XX
XX DT 02-APR-2003 (first entry)
XX XX

DE Lung cancer-associated polynucleotide #196.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;

KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

DR P-PSDB; ABUS6603.

XX Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased

PT expression in lung cancer.

XX Claim 22; Page 335; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated

CC transcript in a cell from a patient, comprising contacting a biological

CC sample from the patient with a polynucleotide that selectively hybridises

CC to a sequence that is at least 80 % identical to a gene that exhibits

CC increased or decreased expression in lung cancer samples. Lung cancer-

CC associated polynucleotides and polypeptides are used for identifying a

CC compound that modulates a lung cancer-associated polypeptide, for

CC inhibiting proliferation of a lung cancer-associated cell to treat lung

CC cancer in a patient and for treating a mammal having lung cancer by

CC administering a modulatory compound identified. The methods are useful

CC for treating lung cancer, such as small cell lung cancer, non-small cell

CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,

CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,

CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and

CC bronchiectasis. The genes, polynucleotides and polypeptides are useful

CC for diagnostic purposes and as targets for screening for therapeutic

CC compounds that modulate lung cancer, such as antibodies. Sequences

CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the

CC invention

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.93 Length: 2053

Score: 44.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x ABX76332 (1-2053)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

Db 373 TTCCTTACCGCAACCAACGAGTGGCGGTG 399

RESULT 13

AAD56197

ID AAD56197 standard; DNA; 2053 BP.

XX

AC AAD56197;

XX

DT 07-AUG-2003 (first entry)

XX

DE Human LRRCAPS DNA #11.

XX

KW Human; p53 pathway; Leucine rich repeat capricious related protein;

KW LRRCAPS; cancer; gene therapy; ds.

XX

OS Homo sapiens.

XX

PN WO2003035831-A2.

XX

PD 01-MAY-2003.

XX

PF 21-OCT-2002; 2002WO-US033540.

XX

PR 22-OCT-2001; 2001US-0338733P.

PR 15-FEB-2002; 2002US-0357600P.

PR 01-MAR-2002; 2002US-0361196P.

XX

PA (EXEL-) EXELIXIS INC.

XX

PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;

PI Francis-Lang H, Friedman L;

XX

DR WPI; 2003-421410/39.

XX

PT Identifying a candidate p53 pathway-modulating agent for treating cancer

PT comprises contacting an assay system comprising a purified leucine rich

PT repeat, capricious related polypeptide or nucleic acid with a test agent.

XX

PS Example 5; Page 73-74; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway

CC modulating agent. The method involves contacting an assay system

CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)

CC polypeptide or nucleic acid or its fragment with a test agent and

CC detecting a test agent-biased activity, where a difference between the

CC test agent-biased activity and the reference activity identifies the test

CC agent as a candidate p53 pathway modulating agent. The method is useful

CC for identifying a candidate p53 pathway-modulating agent for preparing a

CC composition for diagnosing or treating cancer. The invention is useful in

CC gene therapy. The present sequence is human LRRCAPS DNA

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.93 Length: 2053

Score: 44.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x AAD56197 (1-2053)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

Db 373 TTCCTTACCGCAACCAACGAGTGGCGGTG 399

RESULT 14

AAD56200

ID AAD56200 standard; DNA; 2053 BP.

XX

AC AAD56200;

XX

DT 07-AUG-2003 (first entry)

XX

DE Human LRRCAPS DNA #12.
 XX
 KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003035831-A2.
 XX
 PD 01-MAY-2003.
 XX
 PP 21-OCT-2002; 2002WO-US033540.
 PR
 PR 22-OCT-2001; 2001US-0338733P.
 PR
 PR 15-FEB-2002; 2002US-0357600P.
 PR
 PR 01-MAR-2002; 2002US-0361196P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 PI
 XX WPI; 2003-421410/39.
 DR
 XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 XX Disclosure; Page 76-77; 99pp; English.
 XX
 XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. NO.: 9.93 Length: 2053
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x AAD56200 (1-2053)
 Qy 1 PhleuThrGlyAsnGlnleuAlaVal 9
 |||||
 Db 373 TTCCTTACCGGCACACACAGCTGCGCGT 399
 |||||
 RESULT 15
 ADN38721
 ID ADN38721 standard; cDNA; 2053 BP.
 XX
 AC ADN38721;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:39.
 DE
 DE Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;

KW vulnery; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 XX 13-NOV-2002; 2002WO-US036810.
 XX
 PP 13-NOV-2001; 2001US-0350666P.
 PR
 PR 21-NOV-2001; 2001US-0332464P.
 PR
 PR 29-NOV-2001; 2001US-0334393P.
 PR
 PR 03-DEC-2001; 2001US-0335394P.
 PR
 PR 14-DEC-2001; 2001US-0340376P.
 PR
 PR 08-JAN-2002; 2002US-0347211P.
 PR
 PR 10-JAN-2002; 2002US-0347349P.
 PR
 PR 08-FEB-2002; 2002US-0355250P.
 PR
 PR 13-FEB-2002; 2002US-0356714P.
 PR
 PR 20-FEB-2002; 2002US-0359077P.
 PR
 PR 29-MAR-2002; 2002US-0368809P.
 PR
 PR 04-APR-2002; 2002US-0370110P.
 PR
 PR 12-APR-2002; 2002US-0372246P.
 PR
 PR 05-JUN-2002; 2002US-0386614P.
 PR
 PR 16-JUL-2002; 2002US-0396839P.
 PR
 PR 22-JUL-2002; 2002US-039775P.
 PR
 PR 22-JUL-2002; 2002US-0397845P.
 PR
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Afar D, Ariz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 XX WPI; 2003-468649/44.
 DR
 DR P-PSDB; ADN38722.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 XX Claim 8; SEQ ID NO 39; 1385pp; English.
 PS
 XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. NO.: 9.93 Length: 2053
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

US-10-774-176-5 (1-9) x ADN38721 (1-2053)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 373 TTCTTACCGCAACCAAGCTGGCGTG 399

RESULT 16

ADL06473
ID ADL06473 standard; cDNA; 2053 BP.

XX AC ADL06473;

XX DT 20-MAY-2004 (first entry)

XX DE Human tumour-associated antigenic target (TAT) cDNA sequence #53.

XX KW Human; tumour-associated antigenic target; TAT; cell death; tumour;

XX KW cancer; cytostatic; gene; ss.

XX OS Homo sapiens.

XX PN WO2004016225-A2.

XX PD 26-FEB-2004.

XX PF 19-AUG-2003; 2003WO-US025892.

XX PR 19-AUG-2002; 2002US-0404809P.

XX PR 21-AUG-2002; 2002US-0405645P.

XX PR 23-SEP-2002; 2002US-0413192P.

XX PR 15-OCT-2002; 2002US-0419008P.

XX PR 15-NOV-2002; 2002US-0426847P.

XX PR 02-JUL-2003; 2003US-0484959P.

XX PA (GETH) GENENTECH INC.

XX PI Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;

XX PI Spencer SD, Wu TD, Zhang Z;

XX DR WPI; 2004-257144/24.

XX DR P-FSDB; ADL06552.

XX PT New antibody that binds to a tumor-associated antigenic target (TAT)
XX PT polypeptide, useful for preparing a composition for diagnosing or
XX PT treating cancer.
XX PS Claim 1; SEQ ID NO 53; 319pp; English.
XX CC The present invention relates to the isolation of human tumour-associated
XX CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also
XX CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is
XX CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a
XX CC humanised antibody. It is conjugated to a growth inhibitory agent. It is
XX CC produced in bacteria or in CHO cells and induces death of a cell to which
XX CC it binds. The antibody is useful for preparing a composition for
XX CC diagnosing or treating tumours and cancer. The present sequence
XX CC represents a human TAT cDNA sequence of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.93 Length: 2053
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-5 (1-9) x ADL06473 (1-2053)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 373 TTCTTACCGCAACCAAGCTGGCGTG 399

RESULT 17

ADN03961
ID ADN03961 standard; cDNA; 2053 BP.

XX AC ADN03961;

XX DT 01-JUL-2004 (first entry)

XX DE Antipsoriatic cDNA sequence #180.

XX KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.

XX OS Homo sapiens.

XX PN WO2004028479-A2.

XX PD 08-APR-2004.

XX PF 25-SEP-2003; 2003WO-US030907.

XX PR 25-SEP-2002; 2002US-0414006P.

XX PA (GETH) GENENTECH INC.

XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX PI Wu TD;

XX DR WPI; 2004-305105/28.

XX DR P-FSDB; ADN03962.

XX PT New PRO nucleic acid or polypeptide, useful for preparing a
XX PT pharmaceutical composition for diagnosing or treating psoriasis in a
XX PT mammal.

XX PS Claim 1; SEQ ID NO 355; 3069pp; English.

XX CC The invention relates to novel polynucleotide and polypeptides for
XX CC treating psoriasis or a sequence having at least 80% identity to the
XX CC above sequences. The nucleic acid is useful for preparing a composition
XX CC for diagnosing or treating psoriasis in a mammal. This sequence
XX CC corresponds to one of the polynucleotides of the invention.

XX SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.93 Length: 2053
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-5 (1-9) x ADN03961 (1-2053)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

Db 373 TTCTTACCGCAACCAAGCTGGCGTG 399

RESULT 18

ADR25444

ID ADR25444 standard; DNA; 2053 BP.

XX AC ADR25444;

XX DT 21-OCT-2004 (first entry)

XX DE Breast cancer prognosis marker #1305.

XX KW ds; breast cancer; prognosis; gene expression; diagnosis.

XX OS Homo sapiens.

XX PN WO2004065545-A2.

PD 05-AUG-2004.
 XX 15-JAN-2004; 2004WO-US001100.
 XX 15-JAN-2003; 2003US-00342887.
 XX (ROSE-) ROSETTA INPHARMATICS LLC.
 XX (NECA-) NETHERLANDS CANCER INST.
 XX Van't Veer LJ, He Y;
 XX WPI; 2004-593473/57.
 XX
 XX Classifying a breast cancer patient according to prognosis comprises
 XX determining the similarity between the level of expression of each of
 XX five genes in a cell sample taken from patient, to control levels.
 XX
 XX Disclosure; SEQ ID NO 1305; 226pp; English.
 XX
 XX The invention relates to a method of classifying a breast cancer patient
 XX according to prognosis by determining the similarity between the level of
 XX expression of each of five genes for which markers are listed in the
 XX specification, in a cell sample taken from the breast cancer patient, to
 XX control levels of expression for each respective five genes to obtain a
 XX patient similarity value. The methods are useful for classifying a breast
 XX cancer patient according to prognosis. Kits and computer program products
 XX are useful for data analysis using the diagnostic, prognostic and
 XX statistical methods of the invention. This sequence corresponds to a
 XX marker used in the method of the invention.
 XX
 XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 9.93 Length: 2053
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-5 (1-9) x ADR25444 (1-2053)
 Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 373 TTCTTACCGGACACCCAGCTGGCCGTG 399
 RESULT 19
 ACN38510
 ID ACN38510 standard; cDNA; 2053 BP.
 XX
 XX ACN38510;
 XX
 XX 18-NOV-2004 (first entry)
 XX
 XX Tumour-associated antigenic target (TAT) cDNA DNA103471, SEQ ID NO:2070.
 DE
 XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic; gene; ss.
 XX
 XX Homo sapiens.
 XX WO2004030615-A2.
 XX
 XX 15-APR-2004.
 XX
 XX 29-SEP-2003; 2003WO-US028547.
 XX
 XX 02-OCT-2002; 2002US-0414971P.

XX (GETH) GENENTECH INC.
 XX Wu TD, Zhang Z, Zhou Y;
 XX WPI; 2004-347921/32.
 XX P-PSDB; ABM80804.
 XX
 XX New tumor-associated antigenic target polypeptides and nucleic acids,
 XX useful in preparing a medicament for treating or detecting a
 XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 XX prostate cancer or tumor.
 XX
 XX Claim 1; SEQ ID NO 2070; 7273pp; English.
 XX
 XX The invention relates to human tumour-associated antigenic target (TAT)
 XX polypeptides, and their related nucleic acids. The TAT polypeptides are
 XX overexpressed in cancer tissues compared to normal tissues, and may thus
 XX serve as effective targets for the diagnosis and treatment of cancer in
 XX mammals. The invention also relates to nucleic acid and polypeptide
 XX sequences at least 80% identical to the TAT nucleic acids and
 XX polypeptides; expression vectors and host cells comprising a TAT nucleic
 XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
 XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
 XX TAT polypeptide; and methods and compositions for the treatment or
 XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
 XX antibodies, antagonists, binding molecules and compositions are useful
 XX for diagnosing or treating a cell proliferative disorder associated with
 XX increased TAT expression, particularly cancers such as breast cancer,
 XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 XX cancer, pancreatic cancer, cervical cancer, cancers of the central
 XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 XX used as hybridisation probes, in chromosome and gene mapping, in
 XX chromosome identification and in gene therapy. The present sequence
 XX represents a TAT nucleic acid of the invention
 XX
 XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 9.93 Length: 2053
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-5 (1-9) x ACN38510 (1-2053)
 Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 373 TTCTTACCGGACACCCAGCTGGCCGTG 399
 RESULT 20
 ADV35098
 ID ADV35098 standard; cDNA; 2053 BP.
 XX
 XX ADV35098;
 XX
 XX 10-FEB-2005 (first entry)
 XX
 XX Human cDNA of an exemplary efficacy gene for BAD SeqID174.
 DE
 XX human; ss; multi-parameter high throughput screening; MPHTS;
 KW disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
 KW bipolar affective disorder; BAD; autism; Parkinson's;
 KW Alzheimer's disease; neuroleptic; nootropic; antimanic; antidepressant.
 XX
 XX Homo sapiens.
 XX
 XX US2003096264-A1.
 XX
 XX 22-MAY-2003.
 XX

PF 18-JUN-2002; 2002US-00175523.
 XX
 PR 18-JUN-2001; 2001US-0299151P.
 PR 07-SEP-2001; 2001US-0317828P.
 PR 25-SEP-2001; 2001US-0325150P.
 PR 14-NOV-2001; 2001US-0333047P.
 PR 18-JAN-2002; 2002US-0349936P.
 PR 04-MAR-2002; 2002US-0361834P.
 XX
 PA (PSYC-) PSYCHIATRIC GENOMICS INC.
 XX
 PI Altar CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
 PI Palfreyman M, Rajan P;
 XX
 DR WPI; 2004-118903/12.
 XX
 XX Identifying a compound that can treat disease or disorders, such as, a
 PT neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
 PT determining the expression of one or more efficacy genes in a cell
 PT contacted with the test compound.
 XX
 PS Example 6; SEQ ID NO 174; 39pp; English.
 XX
 CC This invention relates to a novel screening method identified as a multi-
 CC parameter high throughput screening (MPTS) assay. Specifically, it
 CC refers to an assay that utilizes the disease signature of a plurality of
 CC specific genes associated with a particular disease, and identifies
 CC differential expression between those cells taken from individuals
 CC affected by that disease and those that are not affected. The present
 CC invention then describes the screening of candidate pharmaceutical
 CC compounds to identify those that have a potential therapeutic benefit for
 CC the treatment of neuropsychiatric and neurodegenerative disorders
 CC including schizophrenia, bipolar affective disorder (BAD) and autism, as
 CC well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
 CC of this invention exhibit various activities including neuroleptic,
 CC neurotropic, antipsychotic and antidepressant. Furthermore, the screening
 CC method used in MPTS will be automated, such that a large number of test
 CC compounds may be rapidly screened with a minimal amount of labour and
 CC effort. This polynucleotide is a human cDNA sequence of a gene that is
 CC differentially expressed in the presence of a therapeutic compound and
 CC represents an exemplary efficacy gene for bipolar affective disorder,
 CC given in an exemplification of the invention.
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9.93 Length: 2053
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-5 (1-9) x ADV35098 (1-2053)
 QY 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 Db 373 TTCCTTACCGGCAACACAGCTGGCGTG 399
 RESULT 21
 AAS87175
 ID AAS87175 standard; cDNA; 2338 BP.
 XX
 AC AAS87175;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 XX DNA encoding novel human diagnostic protein #22979.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.

XX WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 PI Drmanac RT, Liu C, Tang YT;
 XX
 XX P-PSDB; ASG22988.
 DR WPI; 2001-639362/73.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 22979; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (III) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 2338 BP; 519 A; 677 C; 608 G; 534 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 11.6 Length: 2338
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0
 US-10-774-176-5 (1-9) x AAS87175 (1-2338)
 QY 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 Db 630 TTCCTTACCGGCAACACAGCTGGCGTG 656
 RESULT 22
 AAK94253
 ID AAK94253 standard; cDNA; 2359 BP.
 XX
 AC AAK94253;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 XX Human full-length cDNA, SEQ ID NO: 2864.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX OS Homo sapiens.
 XX PN EP1130094-A2.
 XX PD 05-SEP-2001.
 XX PF 07-JUL-2000; 2000EP-00114089.
 XX PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX DR WPI; 2001-524255/58.
 DR P-PSDB; AAK93333.
 XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX PS Claim 8; SEQ ID NO 2864; 1380pp + Sequence Listing; English.
 XX CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 11.7 Length: 2359
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-5 (1-9) x AAK94253 (1-2359)
 QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 712 TTCCTTACCGCAACCACTGGCCGTG 738
 RESULT 23
 ADL30831
 ID ADL30831 standard; cDNA; 2359 BP.
 XX AC ADL30831;
 XX DT 20-MAY-2004 (first entry)
 XX DE Full length human cDNA clone SeqID 2864.
 XX KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.
 OS Homo sapiens.
 XX EP1396543-A2.
 XX PN 10-MAR-2004.
 XX PD

XX 07-JUL-2000; 2003EP-00025638.
 XX PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX DR WPI; 2004-204755/20.
 DR P-PSDB; ADL30832.
 XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 XX PS Example 1; SEQ ID NO 2864; 1340pp; English.
 XX CC This invention relates to a novel primers useful for synthesising full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3',
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.
 XX SQ Sequence 2359 BP; 498 A; 692 C; 634 G; 535 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 11.7 Length: 2359
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-5 (1-9) x ADL30831 (1-2359)
 QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 712 TTCCTTACCGCAACCACTGGCCGTG 738
 RESULT 24
 AAK94254
 ID AAK94254 standard; cDNA; 2361 BP.
 XX AC AAK94254;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human full-length cDNA, SEQ ID NO: 2866.
 XX KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 OS Homo sapiens.
 XX EP1130094-A2.
 XX PN EP1130094-A2.
 XX PD 05-SEP-2001.
 XX PT 07-JUL-2000; 2000EP-00114089.
 XX PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX (HELI-) HELIX RES INST.
 XX PA

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 DR WPI; 2001-524255/58.
 DR P-PSDB; AAM93334.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX
 PS Claim 8; SEQ ID NO 2866; 1380pp + Sequence Listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from BPO
 XX
 SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 11.7 Length: 2361
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-5 (1-9) x AAK94254 (1-2361)
 Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 714 TTCCTTACCGGCAACACGCTGGCCGTG 740
 RESULT 25
 ADI26162
 ID ADI26162 standard; cDNA; 2361 BP.
 XX
 AC ADI26162;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #64.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 PF 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00137505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 PA (ASAH) ASAH KASEI KK.

XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX
 DR WPI; 2004-122214/12.
 DR P-PSDB; ADI26163.
 XX
 PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 PS Claim 4; SEQ ID NO 127; 1368pp; English.
 XX
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 11.7 Length: 2361
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-5 (1-9) x ADI26162 (1-2361)
 Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 Db 714 TTCCTTACCGGCAACACGCTGGCCGTG 740
 RESULT 26
 ADI30833
 ID ADI30833 standard; cDNA; 2361 BP.
 XX
 AC ADI30833;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Full length human cDNA clone SeqID 2866.
 XX
 KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.
 XX
 OS Homo sapiens.
 XX
 PN EPI396543-A2.
 XX
 PD 10-MAR-2004.
 XX
 PF 07-JUL-2000; 2003EP-00025638.
 XX

PR 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183865.
 PR 07-JUL-2000; 2000EP-00114089.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 PA
 XX Ota T, Nishikawa T, Isogai T, Hayaehi K, Iehi S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2004-204755/20.
 DR P-PSDB; ADL30834.
 XX
 XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.
 PT
 XX Example 1; SEQ ID NO 2866; 1340pp; English.
 PS
 XX This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.
 XX
 SQ Sequence 2361 BP; 506 A; 684 C; 638 G; 533 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 11.7 Length: 2361
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-5 (1-9) x ADL30833 (1-2361)
 QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 DB 714 TTCCTTACCGGCAACAGCTGCCGTG 740

RESULT 27
 ABT07721
 ID ABT07721 standard; DNA; 927 BP.
 XX
 AC ABT07721;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Breast cancer-associated gene sequence 29.
 XX
 KW Gene; ds; breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.
 XX
 OS Unidentified.
 XX
 XX WO200259377-A2.
 FN
 XX
 XX 01-AUG-2002.
 PD
 XX
 XX 24-JAN-2002; 2002WO-US002242.
 PF
 XX
 XX 24-JAN-2001; 2001US-0263965P.
 PR
 XX 02-FEB-2001; 2001US-0265928P.
 PR
 XX 09-APR-2001; 2001US-00829472.
 PR
 XX 09-APR-2001; 2001US-0282698P.
 PR
 XX 04-MAY-2001; 2001US-0288590P.
 PR
 XX 29-MAY-2001; 2001US-0294443P.
 PR
 XX (EOSB-) EOS BIOTECHNOLOGY INC. PA

XX Mack DH, Gish KC, Afar D;
 PI
 XX WPI; 2002-583738/62.
 DR N-PSDB; ABJ05564.
 XX
 XX Detecting a breast cancer-associated transcript in a patient's cell,
 PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.
 XX
 XX Claim 9; Page 372; 414pp; English.
 PS
 XX The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
 CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 31.8 Length: 927
 Score: 40.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 90.9% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x ABT07721 (1-927)
 QY 1 PheLeuThrGlyAsnGlnLeuAla 8
 DB 289 TTCCTTACCGGCAACAGCTGCC 312

RESULT 28
 ABX76333
 ID ABX76333 standard; DNA; 927 BP.
 XX
 AC ABX76333;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polynucleotide #197.
 XX
 KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 XX WO200286443-A2.
 FN
 XX
 XX 31-OCT-2002.
 PD
 XX
 XX 18-APR-2002; 2002WO-US012476.
 PF
 XX
 XX 18-APR-2001; 2001US-0284770P.
 PR
 XX 10-MAY-2001; 2001US-0290492P.
 PR
 XX 09-NOV-2001; 2001US-0339245P.
 PR
 XX 13-NOV-2001; 2001US-0350666P.
 PR

PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Aziz N, Murray R;
XX
XX WPI; 2003-093161/08.
DR P-PSDB; ABU56604.
DR
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
XX Claim 22; Page 336; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hyperresistivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 31.8 Length: 927
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.9% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x ABX76333 (1-927)

Qy 1 PheLeuThrGlyAsnGlnLeuAla 8
|||
Db 289 TTCCTTACCGGCAACACGCTGGCC 312

RESULT 29
ADB80503
ID ADB80503 standard; DNA; 927 BP.
XX
XX ADB80503;
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Ovarian cancer-associated transcript #34.
XX
XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
XX post-operative chemotherapy; radiation therapy; tumour prognosis;
XX pre-cancerous lesion detection; ds; gene.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..927
XX /*tag= a
XX

PN WO2002102235-A2.
XX
XX 27-DEC-2002.
XX
XX 18-JUN-2002; 2002WO-US019297.
XX
XX 18-JUN-2001; 2001US-0299234P.
PR 27-AUG-2001; 2001US-0315287P.
PR 05-SEP-2001; 2001US-0317544P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Mack DH, Gish KC;
XX
XX WPI; 2003-167431/16.
DR P-PSDB; ADB80504.
XX
PT Detecting an ovarian cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT polynucleotide that hybridizes to an ovarian cancer gene.
XX
XX Claim 10; Page 297; 332pp; English.
XX
CC The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a
CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC chemotherapeutic tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 31.8 Length: 927
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.9% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-5 (1-9) x ADB80503 (1-927)

Qy 1 PheLeuThrGlyAsnGlnLeuAla 8
|||
Db 289 TTCCTTACCGGCAACACGCTGGCC 312

RESULT 30
ADN38723
ID ADN38723 standard; cDNA; 927 BP.
XX
XX ADN38723;
XX
XX 17-JUN-2004 (first entry)
DT
XX
XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
XX
XX Human; differential expression; cancer; angiogenic disorder;
XX fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
XX inflammatory disease; autoimmune disease;
XX retinal neovascularisation syndrome; scarring; uterine fibroid;
XX detection; diagnosis; prognosis; drug screening; drug targeting;
XX wound healing; contraception; cytostatic; cardiant; immunomodulatory;
XX vulneryary; gene therapy; vaccine; gene; ss.

```
XX OS Homo sapiens.
XX XX WO20003042661-A2.
XX XX 22-MAY-2003.
XX XX 13-NOV-2002; 2002WO-US036810.
XX XX 13-NOV-2001; 2001US-0350666P.
XX XX 21-NOV-2001; 2001US-0332464P.
XX XX 29-NOV-2001; 2001US-0334393P.
XX XX 03-DEC-2001; 2001US-0335394P.
XX XX 14-DEC-2001; 2001US-0340376P.
XX XX 08-JAN-2002; 2002US-0347211P.
XX XX 10-JAN-2002; 2002US-0347349P.
XX XX 08-FEB-2002; 2002US-035250P.
XX XX 13-FEB-2002; 2002US-0356714P.
XX XX 20-FEB-2002; 2002US-0359077P.
XX XX 29-MAR-2002; 2002US-036809P.
XX XX 04-APR-2002; 2002US-0370110P.
XX XX 12-APR-2002; 2002US-0372246P.
XX XX 05-JUN-2002; 2002US-0386614P.
XX XX 16-JUL-2002; 2002US-0396839P.
XX XX 22-JUL-2002; 2002US-039775P.
XX XX 09-SEP-2002; 2002US-0397845P.
XX XX 09-SEP-2002; 2002US-0409450P.
XX XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Heverzi PA;
XX XX Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
XX XX WPI; 2003-468649/44.
XX XX P-PSDB; ADN38724.
XX XX Determining the presence or absence of a pathological cell in a patient,
XX XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX XX a nucleic acid in a biological sample.
XX XX Claim 8; SEQ ID NO 41; 1385pp; English.
XX XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX XX whose expression is upregulated or downregulated in specific cancers or
XX XX other diseases such as angiogenic or fibrotic disorders, and to methods
XX XX of determining the presence or absence of a pathological cell in a
XX XX patient by detecting a nucleic acid at least 80% identical to those of
XX XX the invention or by detecting a polypeptide of the invention. The
XX XX invention also relates to expression vectors and host cells comprising a
XX XX nucleic acid of the invention; antibodies which specifically bind a
XX XX polypeptide of the invention; use of such antibodies for drug targeting;
XX XX and methods of screening for modulators of activity or expression of the
XX XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX XX antibodies and methods are useful for diagnosing, prognosing and treating
XX XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
XX XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX XX neovascularization syndromes, scarring and uterine fibroids. They may
XX XX also be useful in wound healing and in contraception. The present
XX XX sequence represents a nucleic acid sequence of the invention.
XX XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 31.8 Length: 927
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.9% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-5 (1-9) x ADN38723 (1-927)
Oy 1 PheLeuThrGlyAsnGlnLeuAla 8
```

```
Db 289 TTCTTACCGCAACCCAGCTGGCC 312
RESULT 31
ABV99349
ID ABV99349 standard; DNA; 1156 BP.
XX AC ABV99349;
XX XX 27-JAN-2003 (first entry)
XX XX Human NOV8a coding sequence.
XX XX Human; anti-HIV; cytostatic; antidiabetic; antiaethmatic; cachexia; AIDS;
XX XX antinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
XX XX neotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
XX XX antinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
XX XX metabolic disorder; diabetes; obesity; infectious disease; anorexia;
XX XX neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
XX XX immune disorder; haematopoietic disorder; cardiovascular disorder;
XX XX bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
XX XX metabolic syndrome X; wasting disorder; cell differentiation; gene;
XX XX cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX OS Homo sapiens.
XX XX WO200272771-A2.
XX XX 19-SEP-2002.
XX XX 08-MAR-2002; 2002WO-US007288.
XX XX 08-MAR-2001; 2001US-0274101P.
XX XX 08-MAR-2001; 2001US-0274194P.
XX XX 08-MAR-2001; 2001US-0274281P.
XX XX 08-MAR-2001; 2001US-0274322P.
XX XX 09-MAR-2001; 2001US-0274849P.
XX XX 12-MAR-2001; 2001US-0275235P.
XX XX 13-MAR-2001; 2001US-0275578P.
XX XX 13-MAR-2001; 2001US-0275579P.
XX XX 13-MAR-2001; 2001US-0275601P.
XX XX 14-MAR-2001; 2001US-0276000P.
XX XX 16-MAR-2001; 2001US-0276766P.
XX XX 19-MAR-2001; 2001US-0276994P.
XX XX 20-MAR-2001; 2001US-0277239P.
XX XX 20-MAR-2001; 2001US-0277321P.
XX XX 20-MAR-2001; 2001US-0277327P.
XX XX 20-MAR-2001; 2001US-0277338P.
XX XX 21-MAR-2001; 2001US-0277791P.
XX XX 22-MAR-2001; 2001US-0277833P.
XX XX 23-MAR-2001; 2001US-0278152P.
XX XX 26-MAR-2001; 2001US-0278894P.
XX XX 27-MAR-2001; 2001US-0278999P.
XX XX 27-MAR-2001; 2001US-0279036P.
XX XX 28-MAR-2001; 2001US-0279344P.
XX XX 30-MAR-2001; 2001US-0279995P.
XX XX 30-MAR-2001; 2001US-0280233P.
XX XX 02-APR-2001; 2001US-0280802P.
XX XX 02-APR-2001; 2001US-0280822P.
XX XX 02-APR-2001; 2001US-0280900P.
XX XX 04-APR-2001; 2001US-0281194P.
XX XX 13-APR-2001; 2001US-0283675P.
XX XX 30-APR-2001; 2001US-0287424P.
XX XX 03-MAY-2001; 2001US-0288066P.
XX XX 03-MAY-2001; 2001US-0288342P.
XX XX 03-MAY-2001; 2001US-0288528P.
XX XX 15-MAY-2001; 2001US-0291190P.
XX XX 16-MAY-2001; 2001US-0291099P.
XX XX 16-MAY-2001; 2001US-0291240P.
XX XX 30-MAY-2001; 2001US-0294485P.
XX XX 31-MAY-2001; 2001US-0294889P.
XX XX 31-MAY-2001; 2001US-0294899P.
XX XX 18-JUN-2001; 2001US-0299027P.
```

PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 21-NOV-2001; 2001US-0333272P.
 PR 03-DEC-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 04-DEC-2001; 2001US-0338092P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.
 XX

(CURA-) CURAGEN CORP.

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 PI Boldog FI, Li L, Zerhusen BD, Tchervnev VT, Gangolli EA, Vernet CM;
 PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 PI Zhong M;
 XX

DR WPI; 2002-732824/79.

DR P-PSDB; ABP70071.

XX New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 PT disorders, and asthma.

XX Claim 16; Page 114-115; 619pp; English.

XX The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99321-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious diseases, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, haematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods

XX Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	41.2	Length:	1156
Score:	40.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	90.9%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-5 (1-9) x ABV99349 (1-1156)

OY 1 PheLeuThrGlyAsnGlnLeuAla 8
 |||||
 DB 229 TTCTTACCGGCACACGCTGGCC 252

RESULT 32

ADC92179

ID ADC92179 standard; DNA; 972 BP.

XX ADC92179;

XX 01-JAN-2004 (first entry)

DE E. faecium DNA sequence SEQ ID 1806.

XX ds; gene; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.

XX Enterococcus faecium.

XX US6583275-B1.

XX 24-JUN-2003.

XX 30-JUN-1998; 98US-00107532.

XX 02-JUL-1997; 97US-0051571P.

PR 14-MAY-1998; 98US-0085598P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2003-799836/75.

DR P-PSDB; ADC95833.

XX New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.

XX Example 1; SEQ ID NO 1806; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridising to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids are useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium nucleic acids.

XX Sequence 972 BP; 234 A; 166 C; 217 G; 355 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	95.8	Length:	972
Score:	38.00	Matches:	8
Percent Similarity:	88.9%	Conservative:	0
Best Local Similarity:	88.9%	Mismatches:	1
Query Match:	86.4%	Indels:	0
DB:	10	Gaps:	0

US-10-774-176-5 (1-9) x ADC92179 (1-972)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 841 TTTCTGACGGAGGACAGCTCGAGTA 867

RESULT 33
ID AAA27059 standard; DNA; 1281 BP.
XX
AC AAA27059;
XX
DT 22-AUG-2000 (first entry)
XX
DE Mouse 574 tumour-associated antigen gene.
XX
KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
OS Mus musculus.
XX
FN WO200029428-A2.
XX
PD 25-MAY-2000.
XX
PF 18-NOV-1999; 99WO-GB003859.
XX
PR 18-NOV-1999; 98GB-00025303.
PR 27-JAN-1999; 99GB-00001739.
PR 30-JUL-1999; 99GB-00017995.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Carroll MW, Myers KA;
XX
DR WPI; 2000-387735/33.
XX

Tumor associated antigen, 574 capable of eliciting cytotoxic T-lymphocyte
PT response useful in vaccinating against and in treating tumors.
PS Example 2; Page 78; 79pp; English.
XX

The present sequence encodes the mouse 574 tumour-associated antigen
CC (TAA). The TAA 574 is a glycoprotein which is widely expressed in
CC carcinomas but has a highly restricted expression pattern in normal adult
CC tissues. It appears to be strongly correlated to metastasis in colorectal
CC and gastric cancer. 574 antigen may therefore be useful in tumour
CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
CC induced were inoculated with a virus expression vector containing the
CC present sequence. The 574 antigen was shown to be effective at eliciting
CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
CC the antigen and the antigen itself can be used to elicit an immune
CC response, preferably CTL or an antibody response in a subject. The
CC present sequence appears in Genbank at accession number A012160
XX

Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 132 Length: 1281
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 86.4% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-5 (1-9) x AAA27059 (1-1281)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 289 TTCTTACCGGACACAGATACCGTG 315

RESULT 34
ADI26160

ID ADI26160 standard; cDNA; 2557 BP.
XX
AC ADI26160;
XX
DT 22-APR-2004 (first entry)
XX
DE Human cDNA encoding protein that promotes STAT6 activation #63.
XX
SS; gene; human; signal transducer and activator of transcription 6;
KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
XX
OS Homo sapiens.
XX
FN WO2003104277-A2.
XX
PD 18-DEC-2003.
XX
PF 05-JUN-2003; 2003WO-JP007123.
XX
PR 05-JUN-2002; 2002JP-00164257.
PR 06-JUN-2002; 2002US-0385912P.
PR 26-DEC-2002; 2002JP-00377326.
PR 27-DEC-2002; 2002US-0436467P.
PR 15-MAY-2003; 2003JP-00137505.
PR 16-MAY-2003; 2003US-0470836P.
XX
PA (ASAH) ASahi KASEI KK.
XX

Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
WPI; 2004-122214/12.
DR P-PSDB; ADI26161.
XX

New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX

Claim 4; SEQ ID NO 125; 1368pp; English.
PS

The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infections disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.
XX

Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 296 Length: 2557
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1

Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 86.4% Indels: 0
 DB: 12 Gaps: 0

US-10-774-176-5 (1-9) x ADI26160 (1-2557)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 |||||
 Db 844 TTCCTTACCGGCACACGATGACCGTG 870

RESULT 35
 ADI26158
 ID ADI26158 standard; cDNA; 2557 BP.
 XX
 AC ADI26158;
 XX

22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #62.
 XX

ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX

05-JUN-2003; 2003WO-JP007123.
 XX
 PF
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00137505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 PA (ASAH) ASahi KASEI KK.
 XX

Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 WPI; 2004-122214/12.
 DR P-P8DB; ADI26159.
 XX

New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 Claim 4; SEQ ID NO 123; 1368pp; English.
 XX

The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,

CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX

Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;
 SQ

Alignment Scores: 296 Length: 2557
 Pred. No.: 38.00 Matches: 7
 Score: 88.9% Conservative: 1
 Percent Similarity: 77.8% Mismatches: 1
 Best Local Similarity: 86.4% Indels: 0
 Query Match: 12 Gaps: 0
 DB:

US-10-774-176-5 (1-9) x ADI26158 (1-2557)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 |||||
 Db 844 TTCCTTACCGGCACACGATGACCGTG 870

RESULT 36
 ADO35939/c
 ID ADO35939 standard; DNA; 2557 BP.
 XX
 AC ADO35939;
 XX

26-AUG-2004 (first entry)
 XX
 DE Novel mouse gene sequence #612.
 XX

mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
 KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
 KW viral disorder; ds; gene.
 XX
 OS Mus sp.
 XX
 WO2004046310-A2.
 XX
 PD 03-JUN-2004.
 XX

24-OCT-2003; 2003WO-US033948.
 XX
 PR 15-NOV-2002; 2002US-0426916P.
 PR 04-DEC-2002; 2002US-0431158P.
 PR 05-DEC-2002; 2002US-0431445P.
 PR 05-DEC-2002; 2002US-0431606P.
 PR 09-JUN-2003; 2003US-0476621P.
 PR 09-JUN-2003; 2003US-0476632P.
 PR 08-JUL-2003; 2003US-0485217P.
 PR 08-JUL-2003; 2003US-0485359P.
 PR 08-AUG-2003; 2003US-0493332P.
 PR 08-AUG-2003; 2003US-0493356P.
 XX
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX

Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
 WPI; 2004-431966/40.
 XX

New mouse nucleic acid molecules and polypeptides, useful for treating
 PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
 PT disease or thrombosis.
 XX
 Claim 1; SEQ ID NO 612; 263pp; English.
 XX

The invention comprises 744 novel mouse DNA sequences (genes). The DNA
 CC sequences of the invention are useful for treating cancer, psoriasis,
 CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
 CC immune disorders, bacterial disorders and viral disorders. The present
 CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
 CC present DNA sequence is not shown in the specification, but has been

CC retrieved from the WIPO website.

XX
SQ Sequence 2557 BP; 610 A; 794 C; 688 G; 465 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 296 Length: 2557
Score: 39.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 86.4% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-5 (1-9) x ADO35939 (1-2557)

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

DB 57 TTCCTTACCGCAACCAAGATGACCGTG 31

RESULT 37

ACH94847/c

ID ACH94847 standard; DNA; 342 BP.

XX

AC ACH94847;

XX

DT 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polynucleotide seqid 642.

DE

XX Recombinant expression vector; transcription regulatory element;

KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.

XX

OS Klebsiella pneumoniae.

XX

FN US6610836-B1.

XX

PD 26-AUG-2003.

XX

PF 27-JAN-2000; 2000US-00489039.

XX

PR 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PA

XX Breton GL, Osborne M;

XX

XX WPI; 2003-895346/82.

DR P-PSDB; ABO61296.

XX

XX

PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for

preparing a vaccine composition against Klebsiella pneumoniae.

XX

PS Disclosure; SEQ ID NO 642; 932pp; English.

XX

CC The invention describes a new isolated nucleic acid encoding a Klebsiella

pneumoniae polypeptide. Also described are: a recombinant expression

CC vector comprising the nucleic acid, operably linked to a transcription

CC regulatory element; and a cell comprising the recombinant expression

CC vector. The nucleic acid is useful for preparing a vaccine composition

CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella

CC pneumoniae polypeptide of the invention

XX

SQ Sequence 342 BP; 81 A; 88 C; 101 G; 72 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 47.8 Length: 342
Score: 37.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 84.1% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-5 (1-9) x ACH94847 (1-342)

XX

XX

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

DB 171 TTCATCACTGGTGACCACGCTTGCACTA 145

RESULT 38

ABL24406/c

ID ABL24406 standard; DNA; 1553 BP.

XX

AC ABL24406;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 24691.

XX

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ds.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US009231.

XX

PR 23-MAR-2000; 2000US-0191637P.

XX

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE) PE CORP NY.

XX

XX Venter JC, Adams M, Li PWD, Myers EW;

XX

XX WPI; 2001-656860/75.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

XX

PS Claim 1; SEQ ID NO 24691; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1553 BP; 383 A; 332 C; 335 G; 503 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 279 Length: 1553
Score: 37.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 84.1% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-5 (1-9) x ABL24406 (1-1553)

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

DB 1105 TTCITGACAGGAACAAGTTGCTTTA 1079

RESULT 39

ABZ32426/c

ID ABZ32426 standard; DNA; 303 BP.

XX

AC ABZ32426;

XX

AC ABA09091;
 XX 11-JAN-2002 (first entry)
 XX Human secreted protein homologue-encoding cDNA, SEQ ID NO:867.
 DE
 XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157188-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 05-FEB-2001; 2001WO-US003800.
 PP
 XX 03-FEB-2000; 2000US-00496914.
 PR
 XX 27-APR-2000; 2000US-00560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI: 2001-457740/49.
 DR P-PSDB; ABB11847.
 DR
 XX Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 XX
 PS Claim 1; Page 765-766; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC thrombotic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.

CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention
 XX
 SQ Sequence 396 BP; 98 A; 77 C; 96 G; 125 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 95.7 Length: 396
 Score: 36.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 81.8% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-5 (1-9) x ABA09091 (1-396)
 QY 2 LeuThrClyAsnGlnLeuAlaVal 9
 DB 100 ATTACAGGGAACCACTGGCTGTG 123
 RESULT 42
 AAC00706/c
 ID AAC00706 standard; cDNA; 407 BP.
 XX
 AC AAC00706;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 704.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 XX 21-FEB-2000; 2000EP-00200610.
 PF
 PR 26-FEB-1999; 99US-0122487P.
 XX
 PA (GEST) GENSET.
 XX
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 PI WPI: 2000-500381/45.
 DR P-PSDB; AAG00700.
 DR
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 1; SEQ ID NO 704; 71pp + Sequence Listing; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain

CC upstream regulatory sequences and to design expression and secretion
 CC vectors
 CC
 SQ Sequence 407 BP; 124 A; 99 C; 80 G; 104 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 98.8 Length: 407
 Score: 36.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 81.8% Indels: 0
 DB: 3 Gaps: 0

US-10-774-176-5 (1-9) x AAC00706 (1-407)

Oy 2 LeuThrGlyAsnGlnLeuAlaVal 9
 Db 345 ATTACGGGAAACCACTGGCTGTG 322

RESULT 43

ACA48087/C
 ID ACA48087 standard; DNA; 1383 BP.

AC ACA48087;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #29744.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.

OS Streptococcus mutans.

XX WO200277183-A2.

PD 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ASU44217.

XX New antisense nucleic acids, useful for identifying proteins or screening
 FT for homologous nucleic acids required for cellular proliferation to
 FT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 35957; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1383 BP; 381 A; 240 C; 342 G; 420 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 412 Length: 1383
 Score: 36.00 Matches: 6
 Percent Similarity: 88.9% Conservative: 2
 Best Local Similarity: 66.7% Mismatches: 1
 Query Match: 81.8% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x ACA48087 (1-1383)

Oy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

Db 60 TTCCTTACCGGTATCATGCGGATC 34

RESULT 44

ADS11585

ID ADS11585 standard; DNA; 2273 BP.

AC ADS11585;

DT 16-DEC-2004 (first entry)

DE Human therapeutic contig DNA - SEQ ID 1822.

XX antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnary;
 KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
 KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.

OS Homo sapiens.

XX WO2004080148-A2.

XX 23-SEP-2004.

XX 30-SEP-2003; 2003WO-US030720.

XX 02-OCT-2002; 2002US-0416186P.

XX (NUVE-) NUVELO INC.

XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
 PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
 XX WPI; 2004-668857/65.
 DR P-PSDB; ADS12183.

XX New polynucleotide, useful in preparing a composition for diagnosing or
 FT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
 FT aplastic anemia or cancer for promoting wound healing.

XX Example 2; SEQ ID NO 1822; 718pp; English.

CC The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic contig
CC DNA of the invention. The current sequence is not shown explicitly within
CC the specification but can be accessed from the WIPO web-site.

XX SQ Sequence 2273 BP; 527 A; 612 C; 625 G; 509 T; 0 U; 0 Other;

Alignment Scores: 735 Length: 2273
Pred. No.: 36.00 Matches: 7
Score: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 81.8% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-5 (1-9) x ADS11585 (1-2273)

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
|||||
DB 2212 TTTCACCTGGGACAGACTGCACCTT 2238

RESULT 45
AAL06046/c
ID AAL06046 standard; DNA; 3109 BP.

AC AAL06046;

XX 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 8734.

XX Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.

OS Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001339.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 27-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225370P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
XX used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 8734; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a genomic sequence encoding a
XX protein of the invention
XX
XX SQ Sequence 3109 BP; 993 A; 590 C; 526 G; 1000 T; 0 U; 0 Other;

Alignment Scores:
Pred No.: 1.06e+03 Length: 3109
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 81.8% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-5 (1-9) x AAL06046 (1-3109)

Qy 2 LeuThrGlyAsnGlnLeuAlaVal 9
Db 3018 ATTACAGGGAACCACTGGCTGTG 2995

RESULT 46

ABL98611/c
ID ABL98611 standard; DNA; 3109 BP.
XX
AC ABL98611;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding DNA fragment SEQ ID NO: 3263.
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
FN WO200155317-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001329.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.

```
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232402P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0232404P.
PR 14-SEP-2000; 2000US-0232405P.
PR 21-SEP-2000; 2000US-0232423P.
PR 21-SEP-2000; 2000US-0232424P.
PR 25-SEP-2000; 2000US-0232474P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
XX for preventing, diagnosing and/or treating testicular cancer.
XX
XX Disclosure; SEQ ID NO 3263; 766pp; English.
XX
XX The present invention provides the protein and coding sequences of 973
XX human testicular antigens, and fragments of their genomic sequences. The
XX sequences can be used in the treatment of cardiovascular, urinary system,
XX reproductive system, immune, respiratory, neurological and
XX gastrointestinal disorders, infections, and particularly cancer,
XX especially testicular cancers. The present sequence is a DNA encoding a
XX protein fragment of the invention
XX
XX Sequence 3109 BP; 993 A; 590 C; 526 G; 1000 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.06e+03 Length: 3109
Score: 36.00 Matches: 7
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 87.5% Mismatches: 0
Query Match: 81.8% Indels: 0
DB: Gaps: 0
US-10-774-176-5 (1-9) x ABL98611 (1-3109)
QY 2 LeuThrGlyAsnGlnLeuAlaVal 9
Db 3018 ATTACAGGGACCACTGGCTGTG 2995
RESULT 47
ACL28854/C
ID ACL28854 standard; cDNA; 3180 BP.
XX
XX ACL28854;
XX
XX 02-JUN-2005 (first entry)
XX
XX Rice abiotic stress responsive polynucleotide SEQ ID NO:2810.
XX
XX as; abiotic stress tolerance; transgenic plant; plant; cereal;
XX agriculture.
XX
XX Oryza sativa.
XX
XX WO2003008540-A2.
XX
XX 30-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-US019668.
XX
XX 22-JUN-2001; 2001US-0300112P.
XX
XX 24-AUG-2001; 2001US-0314662P.
XX
XX 26-SEP-2001; 2001US-0325277P.
XX
XX 21-NOV-2001; 2001US-0332132P.
```

XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX PA Krepis J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 XX PI Moughamer T, Provart N, Ricke D, Zhu T;
 XX XX MPI, 2003-248011/24.
 XX PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX PT Claim 1; SEQ ID NO 2810; 89pp; English.
 XX CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stress. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 XX SQ Sequence 3180 BP; 1056 A; 596 C; 784 G; 744 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.09e+03 Length: 3180
 Score: 36.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 81.8% Indels: 0
 DB: 11 Gaps: 0
 US-10-774-176-5 (1-9) x ACL28854 (1-3180)
 Qy 1 PheLeuThrGlyAanGlnLeu 7
 Db 2499 TTCTCAGTGGCAACCACTT 2479
 RESULT 48
 ADI45342/c
 ID ADI45342 standard; cDNA; 3180 BP.
 XX AC ADI45342;
 XX DT 22-APR-2004 (first entry)
 XX DE Rice isoprenoid biosynthesis-associated cDNA #137.
 XX KW Rice; isoprenoid biosynthesis; ss; gene; plant; isopentenyl diphosphate;
 KW IPP; dimethylallyl alcohol; DMAPP; short-chain plastid prenyltransferase;
 KW gibberellin; carotenoid; abscisic acid; tocopherol; plastoquinone;
 KW phytylquinone; mevalonate pathway; phytyl; phytyl; phytyl; phytyl;
 KW ubiquinone; monoterpane; sesquiterpene; protein prenylation; chlorophyll;
 KW haeme; Yield.
 XX OS Oryza sativa.
 XX PN US2004010815-A1.
 XX PD 15-JAN-2004.
 XX XX 26-SEP-2002; 2002US-00259194.
 XX PR 26-SEP-2001; 2001US-0325277P.
 PR 04-APR-2002; 2002US-0370620P.
 PR 04-APR-2002; 2002US-0370743P.

XX (LANG/) LANGE B M.
 PA (CHAS/) CHASSEMIAN M.
 PA (BRIG/) BRIGGS S P.
 PA (COOP/) COOPER B.
 PA (GLAZ/) GLAZEBROOK J.
 PA (GOFF/) GOFF S A.
 PA (KATA/) KATAGIRI F.
 PA (KREP/) KREPS J.
 PA (MOUG/) MOUGHAMER T.
 PA (PROV/) PROVART N.
 PA (RICK/) RICKE D.
 PA (ZHUT/) ZHU T.
 XX PI Lange BM, Chasseman M, Briggs SP, Cooper B, Glazebrook J;
 PI Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Ricke D;
 PI Zhu T;
 XX WPI; 2004-090562/09.
 DR P-PSDB; ADI45343.
 XX New isolated polynucleotides and polypeptides associated with isoprenoid
 PT synthesis in plants, useful for producing transgenic plants, for targeted
 PT gene disruption, as well as markers or probes.
 XX Claim 1; SEQ ID NO 273; 117pp; English.
 XX The invention relates to a polynucleotide (or its complement, protein
 CC encoding fragment or reverse complement), comprising a nucleotide
 CC sequence encoding a polypeptide comprising an amino acid sequence
 CC involved in or associated with the biosynthesis of isoprenoids in a rice
 CC plant. Also included are an isolated polypeptide involved in or
 CC associated with the biosynthesis of isoprenoids in a plant, an expression
 CC cassette comprising the polynucleotide, a host cell comprising the
 CC expression cassette, and a transgenic plant comprising the expression
 CC cassette. The polypeptides and polynucleotides include those associated
 CC with the biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl
 CC alcohol (DMAPP), the biosynthesis of short-chain plastid
 CC prenyltransferases, the biosynthesis of gibberellins, the biosynthesis of
 CC carotenoids and/or abscisic acids, the biosynthesis of tocopherols,
 CC plastoquinone and/or phytylquinone biosynthesis, the mevalonate pathway,
 CC phytyl and brassinosteroid metabolism, biosynthesis of ubiquinone,
 CC biosynthesis of monoterpenes and sesquiterpenes, protein prenylation, and
 CC biosynthesis of chlorophyll or haeme. Also disclosed are banana, wheat
 CC and corn homologues of some of the rice polynucleotides. The
 CC polynucleotides are useful for producing transgenic plants, where the
 CC genome is augmented by a nucleic acid molecule of the invention, or in
 CC which the corresponding gene has been disrupted, e.g. to result in a
 CC loss, a decrease or an alteration in the function of the product encoded
 CC by the gene. The plants may also have increased yields and/or produce a
 CC better quality product than the corresponding wild-type plant. The
 CC nucleic acid molecules are useful for targeted gene disruption, as well
 CC as markers and probes. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20040010815. The present sequence
 CC is a Rice isoprenoid biosynthesis-associated cDNA of the invention.
 XX SQ Sequence 3180 BP; 1056 A; 596 C; 784 G; 744 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.09e+03 Length: 3180
 Score: 36.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 81.8% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-5 (1-9) x ADI45342 (1-3180)
 Qy 1 PheLeuThrGlyAanGlnLeu 7
 Db 2499 TTCTCAGTGGCAACCACTT 2479

RESULT 49
AAS85053/c
ID AAS85053 standard; cDNA; 5833 BP.
XX
AC AAS85053;
XX
DT 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #20857.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
FN WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ASG20866.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.
XX
XX Claim 1; SEQ ID NO 20857; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping. (II) is useful as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 5833 BP; 1482 A; 1383 C; 1331 G; 1637 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	2.21e+03	Length:	5833
Score:	36.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	81.8%	Indels:	0
DB:	5	Gaps:	0

US-10-774-176-5 (1-9) x AAS85053 (1-5833)

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 1074 TTCTGCACAGGAATGAAGTCTGTG 1048
RESULT 50
ADA02798
ID ADA02798 standard; DNA; 52754 BP.
XX
AC ADA02798;
XX
XX 06-NOV-2003 (first entry)
XX
DE Human TNFSE11 carcinoma associated gene, SEQ ID NO:1316.
XX
KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast; prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening; gene; da.
XX
XX Homo sapiens.
XX
XX WO2003057146-A2.
XX
XX 17-JUL-2003.
XX
XX 26-DEC-2002; 2002WO-US041414.
XX
XX 26-DEC-2001; 2001US-00035832.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI; 2003-587068/55.
XX
XX New recombinant nucleic acid encoding carcinoma associated protein, useful for preparing compositions for treating carcinomas.
XX
XX Claim 1; SEQ ID NO 1316; 245pp; English.

XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 52754 BP; 15093 A; 10533 C; 11422 G; 15706 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	2.88e+04	Length:	52754
Score:	36.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	81.8%	Indels:	0
DB:	9	Gaps:	0

US-10-774-176-5 (1-9) x ADA02798 (1-52754)

Qy 1 PheLeuThrGlyAsnGlnLeu 7
|||||
Db 16140 TTCTTACAGGCATCAGCTA 16160

Search completed: April 25, 2006, 12:33:06
Job time : 328.3 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 25, 2006, 10:37:32 ; Search time 2986.7 Seconds
(without alignments)
171.290 Million cell updates/sec

Title: US-10-774-176-5

Perfect score: 44

Sequence: 1 PLTGNQLAV 9

Scoring table:

BLASTSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Dapop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB/spool/US10774176/runat_24042006_165114_19197/app_query.fasta_1
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blastsum62 -TRANS=human40.cdi -LIST=1000
-DOALIGN=200 -THR SCORES=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss04
-USBR=US10774176 @CGEN 1 1 6765 @runat_24042006_165114_19197 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.sts.*

11: gb.sy.*

12: gb.un.*

13: gb.vi.*

14: gb.htg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	1260	6	AX467373 Sequence
2	44	100.0	1260	6	AX821533 Sequence
3	44	100.0	1260	6	AX821548 Sequence

4	44	100.0	1263	6	BD249731
5	44	100.0	1263	6	AX025011
6	44	100.0	1263	6	AX149553
7	44	100.0	1263	6	AX316086
8	44	100.0	1263	6	AX467371
9	44	100.0	2053	8	CQ731678
10	44	100.0	2053	8	HS5T40A
11	44	100.0	2359	6	BD127282
12	44	100.0	2359	6	CQ782724
13	44	100.0	2359	8	AK074786
14	44	100.0	2361	6	BD127283
15	44	100.0	2361	6	CQ782726
16	44	100.0	2361	6	AX961916
17	44	100.0	2361	8	AK074790
18	44	100.0	2379	8	BC037161
19	44	100.0	2714	8	AB168308
20	44	100.0	5551	8	HS4012159
21	44	100.0	121909	8	HSJ492P14
22	41	93.2	88345	14	CR854913
23	41	93.2	168825	14	CR854985
24	40	90.9	927	6	AX829164
25	40	90.9	110000	15	AF008208
26	40	90.9	136267	15	AP005756
27	39	88.6	60199	8	AC083883
28	39	88.6	209688	14	AC160293
29	39	88.6	222692	14	AC126962
30	39	88.6	228883	9	AC127173
31	39	88.6	230689	9	AC159326
32	39	88.6	234843	14	AC106580
33	39	88.6	259795	14	AC128267
34	38	86.4	365	10	AL807132
35	38	86.4	972	6	AR347195
36	38	86.4	1281	6	BD249732
37	38	86.4	1281	6	AX025012
38	38	86.4	1281	6	AX316087
39	38	86.4	2333	9	AF063939
40	38	86.4	2361	9	BC087011
41	38	86.4	2423	9	BC058198
42	38	86.4	2557	6	AX961912
43	38	86.4	2557	6	AX961914
44	38	86.4	7942	9	MMU012160
45	38	86.4	15951	1	AE001073
46	38	86.4	108767	15	AC011809
47	38	86.4	146219	14	CR762469
48	38	86.4	155001	14	AC140093
49	38	86.4	167046	9	AC158516
50	38	86.4	168567	9	AC160108
51	38	86.4	171024	14	AC149250
52	38	86.4	171039	14	AC149102
53	38	86.4	176600	9	AC113508
54	38	86.4	203518	14	AC149169
55	38	86.4	208083	9	AC133514
56	38	86.4	210237	14	AC128294
57	38	86.4	211770	14	AC159578
58	38	86.4	218774	14	AC079182
59	38	86.4	231023	14	AC102598
60	38	86.4	231177	14	AC158219
61	38	86.4	234652	9	AC140364
62	38	86.4	239076	14	AC108962
63	38	86.4	280895	14	AC161901
64	38	86.4	289640	14	AC150758
65	37	84.1	257	6	CQ683469
66	37	84.1	342	6	AR383913
67	37	84.1	881	6	AR495929
68	37	84.1	881	6	AR511211
69	37	84.1	1388	4	BOVALRDD
70	37	84.1	1553	6	CQ606092
71	37	84.1	1611	2	AF303112
72	37	84.1	1867	2	BT022877
73	37	84.1	1909	2	BT022814
74	37	84.1	1914	2	BT022853
75	37	84.1	15502	14	AC013227
76	37	84.1	44876	14	AC100114

BD249731	Polypepti
AX025011	Sequence
AX149553	Sequence
AX316086	Sequence
AX467371	Sequence
CQ731678	Sequence
HS5T40A	Sequence
BD127282	Primer fo
CQ782724	Sequence
AK074786	Homo sapi
BD127283	Primer fo
CQ782726	Sequence
AX961916	Sequence
AK074790	Homo sapi
BC037161	Homo sapi
AB168308	Macaca fa
AJ012159	Homo sapi
AL121977	Human DNA
Continuation (4 of	
CR854985	Danio rer
AX829164	Sequence
Continuation (74 o	
AP005756	Oryza sat
AC083883	Homo sapi
AC160293	Bos tauru
AC126962	Rattus no
AC127173	Mus muscu
AC159326	Mus muscu
AC106580	Rattus no
AC128267	Rattus no
AL807132	Arabidops
AR347195	Sequence
BD249732	Polypepti
AX025012	Sequence
AX316087	Sequence
AF063939	Rattus no
BC087011	Rattus no
BC058198	Mus muscu
AX961912	Sequence
AX961914	Sequence
AJ012160	Mus muscu
AE001073	Archaeogl
AC011809	Arabidops
CR762469	Danio rer
AC140093	Bos tauru
AC158516	Mus muscu
AC160108	Mus muscu
AC149250	Papio anu
AC149102	Papio anu
AC113508	Mus muscu
AC149169	Papio anu
AC133514	Mus muscu
AC128294	Rattus no
AC159578	Papio anu
AC079182	Mus muscu
AC102598	Mus muscu
AC158219	Mus muscu
AC140364	Mus muscu
AC106962	Rattus no
AC161901	Bos tauru
AC150758	Bos tauru
CQ683469	Sequence
AR383913	Sequence
AR495929	Sequence
AR511211	Sequence
MS9756	Bovine lens
CQ606092	Sequence
AF303112	Physarum
BT022877	Drosophil
BT022814	Drosophi
BT022853	Drosophi
AC013227	Dr.
AC100114	

C 77	37	84.1	46765	8	AL354863	AL354863 Human DNA	150	36	81.8	151405	14	AC011172	AC011172 Homo sapi
C 78	37	84.1	50182	14	AC113856	Continuation (4 of	151	36	81.8	153248	14	AC112500	AC112500 Homo sapi
C 79	37	84.1	73521	14	AC139545	AC139545 Homo sapi	C 152	36	81.8	154641	5	BX571701	BX571701 Zebrafish
C 80	37	84.1	91000	8	H93706G20	AL121591 Human DNA	C 153	36	81.8	156532	5	CR450737	CR450737 Zebrafish
C 81	37	84.1	123576	8	AC015969	AC015969 Homo sapi	C 154	36	81.8	158212	5	BX649637	BX649637 Zebrafish
C 82	37	84.1	165588	14	AC156181	AC156181 Bos tauru	C 155	36	81.8	159691	8	AC025160	AC025160 Homo sapi
C 83	37	84.1	166035	9	AC123929	AC123929 Mus muscu	C 156	36	81.8	159711	5	BX322551	BX322551 Zebrafish
C 84	37	84.1	166518	8	AC007431	AC007431 Homo sapi	C 157	36	81.8	160055	14	AC144744	AC144744 Danio rer
C 85	37	84.1	174631	2	AC010698	AC010698 Drosophil	C 158	36	81.8	163204	14	AL356749	AL356749 Homo sapi
C 86	37	84.1	175375	14	AC162073	AC162073 Bos tauru	C 159	36	81.8	163379	14	CR450714	CR450714 Danio rer
C 87	37	84.1	178058	8	AC009139	AC009139 Homo sapi	C 160	36	81.8	163806	8	AC113349	AC113349 Homo sapi
C 88	37	84.1	182030	14	AC162765	AC162765 Rhinoloph	C 161	36	81.8	165089	5	BX004845	BX004845 Zebrafish
C 89	37	84.1	183216	14	AL161726	AL161726 Homo sapi	C 162	36	81.8	165350	14	AC134958	AC134958 Rattus no
C 90	37	84.1	191674	4	CR536601	CR536601 Platypus	C 163	36	81.8	166576	5	BX005141	BX005141 Zebrafish
C 91	37	84.1	199113	8	AL445645	AL445645 Human DNA	C 164	36	81.8	172893	14	AC024596	AC024596 Homo sapi
C 92	37	84.1	202807	9	AC160131	AC160131 Mus muscu	C 165	36	81.8	174208	14	AC015935	AC015935 Homo sapi
C 93	37	84.1	206466	8	AC010548	AC010548 Homo sapi	C 166	36	81.8	175296	9	AC155231	AC155231 Mus muscu
C 94	37	84.1	214241	14	AC157426	AC157426 Bos tauru	C 167	36	81.8	176104	14	AC016193	AC016193 Homo sapi
C 95	37	84.1	215236	9	AC162181	AC162181 Mus muscu	C 168	36	81.8	178918	9	AC120159	AC120159 Mus muscu
C 96	37	84.1	216030	14	AC152304	AC152304 Bos tauru	C 169	36	81.8	180999	14	AC150610	AC150610 Callithri
C 97	37	84.1	221664	14	AC151200	AC151200 Bos tauru	C 170	36	81.8	181931	14	AC133025	AC133025 Rattus no
C 98	37	84.1	222741	9	AC140985	AC140985 Mus muscu	C 171	36	81.8	182483	5	CR853297	CR853297 Zebrafish
C 99	37	84.1	243289	14	AC128764	AC128764 Rattus no	C 172	36	81.8	182432	5	BX957239	BX957239 Zebrafish
C 100	37	84.1	252738	14	AC159838	AC159838 Bos tauru	C 173	36	81.8	183192	9	AY016021	AY016021 Mus muscu
C 101	37	84.1	276132	2	AB003543	AB003543 Drosophil	C 174	36	81.8	184872	9	AL672243	AL672243 Mouse DNA
C 102	37	84.1	287233	14	AC110639	AC110639 Rattus no	C 175	36	81.8	184906	14	BX914208	BX914208 Danio rer
C 103	36	81.8	303	6	AX489413	AX489413 Sequence	C 176	36	81.8	186609	8	AC083872	AC083872 Homo sapi
C 104	36	81.8	332	6	BD072161	BD072161 Secreted	C 177	36	81.8	187098	14	AC015984	AC015984 Homo sapi
C 105	36	81.8	407	6	BD024451	BD024451 Sequence	C 178	36	81.8	187532	8	AC006333	AC006333 Bos tauru
C 106	36	81.8	407	6	AX884841	AX884841 Sequence	C 179	36	81.8	188317	15	OSJN00245	OSJN00245 Oryza sat
C 107	36	81.8	510	4	AB098903	AB098903 Bos tauru	C 180	36	81.8	188765	8	AC118755	AC118755 Homo sapi
C 108	36	81.8	578	10	G49021	G49021 SHGC-78189	C 181	36	81.8	189338	14	CR847789	CR847789 Danio rer
C 109	36	81.8	698	10	BV625221	BV625221 S216P6046	C 182	36	81.8	189301	9	AC132225	AC132225 Mus muscu
C 110	36	81.8	1000	10	CN80618Y	AL400100 T3 end of	C 183	36	81.8	189684	14	AC152311	AC152311 Bos tauru
C 111	36	81.8	1119	5	AY874346	AY874346 Xenopus f	C 184	36	81.8	190078	14	AC108895	AC108895 Bos tauru
C 112	36	81.8	1119	6	AR556739	AR556739 Sequence	C 185	36	81.8	192197	8	AC087501	AC087501 Homo sapi
C 113	36	81.8	1135	5	AY874348	AY874348 Xenopus b	C 186	36	81.8	194255	14	CR925728	CR925728 Danio rer
C 114	36	81.8	1141	5	AY874347	AY874347 Xenopus r	C 187	36	81.8	195584	14	AC147540	AC147540 Pan trogl
C 115	36	81.8	1141	5	AY874350	AY874350 Xenopus l	C 188	36	81.8	195900	14	AC021164	AC021164 Homo sapi
C 116	36	81.8	1142	5	AY874344	AY874344 Xenopus w	C 189	36	81.8	196260	14	CR931819	CR931819 Danio rer
C 117	36	81.8	1156	4	BOVALDRED	M31463 Bovine aldo	C 190	36	81.8	197243	14	AC134490	AC134490 Rattus no
C 118	36	81.8	1187	4	S54973	S54973 20 alpha-hy	C 191	36	81.8	200724	8	AL139382	AL139382 Human DNA
C 119	36	81.8	1207	15	AY138789	AY138789 Rhizopus	C 192	36	81.8	202116	14	AC142551	AC142551 Danio rer
C 120	36	81.8	1354	4	BT021058	BT021058 Bos tauru	C 193	36	81.8	203187	9	AC124168	AC124168 Mus muscu
C 121	36	81.8	1400	4	BOVALRDC	MS9755 Bovine lens	C 194	36	81.8	204727	9	AC147020	AC147020 Mus muscu
C 122	36	81.8	2359	4	BOVALRDB	MS9754 Bovine lens	C 195	36	81.8	204730	5	BX890575	BX890575 Zebrafish
C 123	36	81.8	4049	15	AK066916	AK066916 Oryza sat	C 196	36	81.8	206301	14	BX927312	BX927312 Danio rer
C 124	36	81.8	10085	1	AE006222	AE006222 Pasteurel	C 197	36	81.8	207486	8	AL356583	AL356583 Human DNA
C 125	36	81.8	12027	1	AE0014971	AE0014971 Streptoco	C 198	36	81.8	208409	14	AC117895	AC117895 Rattus no
C 126	36	81.8	15722	2	AF003132	AF003132 Caenorhab	C 199	36	81.8	211987	14	CR751564	CR751564 Danio rer
C 127	36	81.8	20653	1	AE008836	AE008836 Salmonell	C 200	36	81.8	212454	14	CT009613	CT009613 Mus muscu
C 128	36	81.8	34978	2	CBRG41N04	AC084590 Caenorhab	C 201	36	81.8	215111	8	AC010255	AC010255 Homo sapi
C 129	36	81.8	52754	6	AX695689	AX695689 Sequence	C 202	36	81.8	215819	14	AC099288	AC099288 Rattus no
C 130	36	81.8	68702	8	AL445193	AL445193 Human DNA	C 203	36	81.8	216194	9	AC096051	AC096051 Rattus no
C 131	36	81.8	72576	9	EX005243	EX005243 Mouse DNA	C 204	36	81.8	217604	14	AC152637	AC152637 Bos tauru
C 132	36	81.8	85624	8	AL139393	AL139393 Human DNA	C 205	36	81.8	218039	5	CR354584	CR354584 Zebrafish
C 133	36	81.8	108082	8	AC110759	AC110759 Homo sapi	C 206	36	81.8	218075	14	AC161920	AC161920 Bos tauru
C 134	36	81.8	103419	14	AL139221	AL139221 Homo sapi	C 207	36	81.8	220373	14	AC105509	AC105509 Rattus no
C 135	36	81.8	110000	1	CP000026	Continuation (30 o	C 208	36	81.8	223345	14	AC123324	AC123324 Rattus no
C 136	36	81.8	110000	14	AC111775	Continuation (2 of	C 209	36	81.8	224843	14	AC095744	AC095744 Rattus no
C 137	36	81.8	110000	15	AP008209	Continuation (78 o	C 210	36	81.8	226768	14	AC109788	AC109788 Rattus no
C 138	36	81.8	110000	15	AP008209	Continuation (190	C 211	36	81.8	230050	1	AL627277	AL627277 Salmonell
C 139	36	81.8	118812	9	AL929446	AL929446 Mouse DNA	C 212	36	81.8	231389	14	AC098169	AC098169 Rattus no
C 140	36	81.8	121145	5	BX927123	BX927123 Zebrafish	C 213	36	81.8	235516	14	AC163072	AC163072 Bos tauru
C 141	36	81.8	129992	8	AL354897	AL354897 Human DNA	C 214	36	81.8	239385	14	AC109686	AC109686 Rattus no
C 142	36	81.8	131217	13	AY528864	AY528864 Macaca mu	C 215	36	81.8	240971	14	AC120956	AC120956 Rattus no
C 143	36	81.8	137453	5	BX927165	BX927165 Zebrafish	C 216	36	81.8	241602	14	AC112881	AC112881 Rattus no
C 144	36	81.8	143769	8	AC004914	AC004914 Homo sapi	C 217	36	81.8	249450	14	AC108355	AC108355 Rattus no
C 145	36	81.8	146930	15	AC118672	AC118672 Genomic s	C 218	36	81.8	250600	14	AC136646	AC136646 Rattus no
C 146	36	81.8	149866	8	AC025627	AC025627 Homo sapi	C 219	36	81.8	250614	14	AC095517	AC095517 Rattus no
C 147	36	81.8	149948	14	AC068852	AC068852 Homo sapi	C 220	36	81.8	251322	14	AC157234	AC157234 Bos tauru
C 148	36	81.8	151144	14	CR848724	CR848724 Danio rer	C 221	36	81.8	256046	14	AC096263	AC096263 Rattus no
C 149	36	81.8	151326	14	AC020864	AC020864 Mus muscu	C 222	36	81.8	256791	9	AC113961	AC113961 Mus muscu

c 223	36	81.8	256927	14	AC099357	Rattus no	AC010137	8	AC010137	Homo sapi
c 224	36	81.8	257608	14	AC153408	Bos tauru	AC108026	8	AC108026	Homo sapi
c 225	36	81.8	257896	14	AC162472	Bos tauru	AC151789	4	AC151789	Ornithorh
c 226	36	81.8	268828	14	AC091513	Rattus no	AL356268	8	AL356268	Human DNA
c 227	36	81.8	281185	14	AC114528	Rattus no	AL101177	8	AL101177	Human DNA
c 228	36	81.8	301311	1	AE016843	Salmonello	AC097011	15	AC097011	Sus scrofa
c 229	36	81.8	306803	1	AB017161	Prochlorella	AC135644	15	AC135644	Oryza sat
c 230	35	79.5	332	10	AB140600	Homo sapi	AC166742	14	AC166742	Glycine m
c 231	35	79.5	352	10	BV332298	Homo sapi	AC135814	8	AC135814	Homo sapi
c 232	35	79.5	359	10	G60804	SHGC-82848	AC108110	8	AC108110	Homo sapi
c 233	35	79.5	364	10	BV096641	RPAMSE90	AC100946	8	AC100946	Homo sapi
c 234	35	79.5	367	6	BD025835	Sequence	AC148378	14	AC148378	Sorex ara
c 235	35	79.5	367	6	AX886225	Sequence	282215	8	282215	Human DNA s
c 236	35	79.5	368	6	CQ395079	Sequence	AC11222	14	AC11222	Homo sapi
c 237	35	79.5	568	6	CQ401421	Sequence	CR387991	14	CR387991	Danio rer
c 238	35	79.5	581	10	BV231334	S233P6373	AC010361	8	AC010361	Homo sapi
c 239	35	79.5	611	10	BV310461	S233P6370	AC109454	8	AC109454	Homo sapi
c 240	35	79.5	624	10	BV020151	S212P6864	AC115848	8	AC115848	Mus muscu
c 241	35	79.5	650	10	BV392987	S243P6205	AC009650	8	AC009650	Homo sapi
c 242	35	79.5	687	8	HUMYX29D10	Homo sapi	BS000055	8	BS000055	Pan trogl
c 243	35	79.5	720	6	CQ407805	Sequence	AC135190	15	AC135190	Oryza sat
c 244	35	79.5	749	6	CQ729454	Sequence	AL021578	8	AL021578	Human DNA
c 245	35	79.5	876	6	CQ740211	Sequence	AC069440	8	AC069440	Homo sapi
c 246	35	79.5	1066	6	CQ413253	Sequence	CS086351	6	CS086351	Sequence
c 247	35	79.5	1067	8	BC022807	Homo sapi	AC021926	14	AC021926	Homo sapi
c 248	35	79.5	1069	6	BD225549	Human pro	AC147764	14	AC147764	Dasytus n
c 249	35	79.5	1071	6	BD231893	Bone marr	AC110513	9	AC110513	Mus muscu
c 250	35	79.5	1082	6	CQ834268	Sequence	AC013361	8	AC013361	Homo sapi
c 251	35	79.5	1109	6	CQ798108	Sequence	AC154085	14	AC154085	Rhinoloph
c 252	35	79.5	1109	6	CS033592	Sequence	AC140195	9	AC140195	Mus muscu
c 253	35	79.5	1109	6	CS036648	Sequence	AC153310	14	AC153310	Otolemur
c 254	35	79.5	1109	6	CS042544	Sequence	AC026279	8	AC026279	Homo sapi
c 255	35	79.5	1109	6	CS045600	Sequence	AC027609	8	AC027609	Homo sapi
c 256	35	79.5	1109	8	AF157321	Homo sapi	AC087525	8	AC087525	Homo sapi
c 257	35	79.5	1223	8	AF274940	Homo sapi	AC100781	8	AC100781	Homo sapi
c 258	35	79.5	1413	6	AX802197	Sequence	AC007777	14	AC007777	Homo sapi
c 259	35	79.5	1920	8	AK222953	Homo sapi	AC121101	9	AC121101	Mus muscu
c 260	35	79.5	2749	8	AB004818	Homo sapi	AC124392	8	AC124392	Mus muscu
c 261	35	79.5	3646	9	AB004818	Musculu	CR352341	14	CR352341	Danio rer
c 262	35	79.5	11966	1	AB002152	Ureaplasma	AC159189	8	AC159189	Mus muscu
c 263	35	79.5	28065	8	AC005354	Homo sapi	BX936352	14	BX936352	Danio rer
c 264	35	79.5	31929	8	HS113988	Human DNA s	AC155786	14	AC155786	Papio anu
c 265	35	79.5	36676	8	HS9858B16	Arabidops	BX005154	14	BX005154	Zebrafish
c 266	35	79.5	38089	15	AB028605	Arabidops	AC121941	9	AC121941	Mus muscu
c 267	35	79.5	41024	14	AC145648	Homo sapi	AL391517	8	AL391517	Human chr
c 268	35	79.5	42165	14	EX004762	Continuation (5 of	AC119876	8	AC119876	Mus muscu
c 269	35	79.5	42973	14	AC145659	Homo sapi	AC123049	8	AC123049	Mus muscu
c 270	35	79.5	43242	8	AC145647	Homo sapi	AC024612	8	AC024612	Homo sapi
c 271	35	79.5	51246	14	AC101011	Mus muscu	AC162858	8	AC162858	Mus muscu
c 272	35	79.5	55007	8	AC142320	Pan trogl	AC023404	14	AC023404	Homo sapi
c 273	35	79.5	56753	14	AL954766	Continuation (4 of	AC007783	8	AC007783	Homo sapi
c 274	35	79.5	59507	14	AC126331	Homo sapi	AC124908	8	AC124908	Equus cab
c 275	35	79.5	62542	14	AC015868	Homo sapi	AC093275	8	AC093275	Homo sapi
c 276	35	79.5	63196	14	AP007765	Lotus cor	AC092644	8	AC092644	Homo sapi
c 277	35	79.5	65321	14	AC068736	Homo sapi	AC018588	8	AC018588	Homo sapi
c 278	35	79.5	66722	14	AC104359	Homo sapi	AC023591	8	AC023591	Homo sapi
c 279	35	79.5	75607	14	AC101635	Mus muscu	AC162118	14	AC162118	Cercopit
c 280	35	79.5	83698	6	AX540653	Sequence	AC022007	8	AC022007	Homo sapi
c 281	35	79.5	83698	15	AB010072	Arabidops	AF083031	15	AF083031	Guillardi
c 282	35	79.5	91073	14	AC011426	Homo sapi	AC122402	9	AC122402	Mus muscu
c 283	35	79.5	91501	14	AC024344	Homo sapi	AC147375	9	AC147375	Mus muscu
c 284	35	79.5	101335	9	AL928822	Mus muscu	AC021099	14	AC021099	Homo sapi
c 285	35	79.5	101406	14	AP008093	Lotus cor	AC087294	8	AC087294	Homo sapi
c 286	35	79.5	101765	8	AL136306	Human DNA	CR354558	8	CR354558	Zebrafish
c 287	35	79.5	110000	1	AB016853	Continuation (59 o	AC107884	8	AC107884	Homo sapi
c 288	35	79.5	110000	2	AB003788	Continuation (2 of	AC128880	14	AC128880	Mus muscu
c 289	35	79.5	110000	14	EX004762	Mus muscu	AC113799	14	AC113799	Rattus no
c 290	35	79.5	110000	14	EX936368	Continuation (3 of	AC020780	14	AC020780	Homo sapi
c 291	35	79.5	110000	15	AP008217	Continuation (269	AC021805	14	AC021805	Homo sapi
c 292	35	79.5	110000	15	AE017341	Continuation (18 o	AL714021	9	AL714021	Mouse DNA
c 293	35	79.5	110261	9	AL807401	Mouse DNA	AC133077	9	AC133077	Mus muscu
c 294	35	79.5	114668	14	AP007632	Lotus cor	AC109244	8	AC109244	Mus muscu
c 295	35	79.5	120282	8	AL136366	Human DNA	AL807805	9	AL807805	Mouse DNA

369	35	79.5 183533	8	BS000021	BS000021 Pan trogl	c 442	35	79.5 258650	1	AL596171	AL596171 Listeria
370	35	79.5 187737	8	AC007297	AC007297 Homo sapi	c 443	35	79.5 260300	14	AC154149	AC154149 Pongo pyg
371	35	79.5 189229	14	AC130132	AC130132 Rattus no	c 444	35	79.5 261369	14	AC106158	AC106158 Rattus no
372	35	79.5 189369	14	AC115650	AC115650 Rattus no	c 445	35	79.5 261522	14	AC094870	AC094870 Rattus no
c 373	35	79.5 190264	14	AC161624	AC161624 Pan trogl	c 446	35	79.5 264437	14	AC117151	AC117151 Rattus no
374	35	79.5 190843	14	AC148362	AC148362 Gaasterost	447	35	79.5 272124	14	AC129654	AC129654 Rattus no
c 375	35	79.5 190936	5	BX548158	BX548158 Zebrafish	c 448	35	79.5 272155	14	AC165426	AC165426 Mus muscu
c 376	35	79.5 192167	9	AC166325	AC166325 Mus muscu	c 449	35	79.5 279353	14	AC120587	AC120587 Rattus no
377	35	79.5 192437	5	BX322665	BX322665 Zebrafish	c 450	35	79.5 290164	14	AC158499	AC158499 Mus muscu
c 378	35	79.5 193214	8	AC069461	AC069461 Homo sapi	c 451	35	79.5 291417	14	AC128252	AC128252 Rattus no
379	35	79.5 195222	2	AC009206	AC009206 Drosophil	452	35	79.5 300953	14	AC162349	AC162349 Bos tauru
380	35	79.5 196457	14	AC163345	AC163345 Mus muscu	453	35	79.5 316900	1	TACID3	AL445065 Thermopila
381	35	79.5 197371	5	BX294000	BX294000 Zebrafish	454	35	79.5 335799	14	BX005139	BX005139 Mus muscu
382	35	79.5 198050	1	AL646061	AL646061 Ralstonia	455	35	79.5 337101	8	HSKSRPXR	AL121578 Homo sapi
383	35	79.5 198546	14	AC069337	AC069337 Homo sapi	456	35	79.5 340000	8	AF001677	AP001677 Homo sapi
c 384	35	79.5 199682	8	BS000620	BS000620 Pan trogl	457	35	79.5 340000	8	HS21C010	AL163210 Homo sapi
c 385	35	79.5 199863	8	AC142325	AC142325 Pan trogl	458	35	79.5 349980	6	CQ870482	CQ870482 Sequence
386	35	79.5 200250	9	AL731697	AL731697 Mouse DNA	c 459	35	79.5 349980	6	AX417047	AX417047 Sequence
c 387	35	79.5 200426	9	AC139759	AC139759 Mus muscu	c 460	34	77.3 177	6	CQ056547	CQ056547 Sequence
c 388	35	79.5 201606	14	AC142521	AC142521 Rattus no	c 461	34	77.3 177	6	CQ058410	CQ058410 Sequence
c 389	35	79.5 201667	14	AC163492	AC163492 Mus muscu	c 462	34	77.3 177	6	CQ075819	CQ075819 Sequence
c 390	35	79.5 202689	9	AL672232	AL672232 Mouse DNA	c 463	34	77.3 177	6	CQ077757	CQ077757 Sequence
c 391	35	79.5 204420	14	AC162624	AC162624 Bos tauru	c 464	34	77.3 177	6	CQ106802	CQ106802 Sequence
392	35	79.5 205302	14	AC018885	AC018885 Homo sapi	c 465	34	77.3 177	6	CQ108762	CQ108762 Sequence
393	35	79.5 207620	8	AP001331	AP001331 Homo sapi	c 466	34	77.3 177	6	CQ116621	CQ116621 Sequence
394	35	79.5 207708	8	AC018808	AC018808 Homo sapi	c 467	34	77.3 177	6	CQ145455	CQ145455 Sequence
395	35	79.5 208183	8	AC144428	AC144428 Pan trogl	c 468	34	77.3 177	6	CQ147394	CQ147394 Sequence
c 396	35	79.5 210175	9	AC128668	AC128668 Mus muscu	c 469	34	77.3 177	6	CQ180898	CQ180898 Sequence
397	35	79.5 213050	1	AL646079	AL646079 Ralstonia	c 470	34	77.3 177	6	CQ182767	CQ182767 Sequence
c 398	35	79.5 213251	6	AX413015	AX413015 Sequence	c 471	34	77.3 177	6	CQ205269	CQ205269 Sequence
c 399	35	79.5 213412	9	AC126044	AC126044 Mus muscu	c 472	34	77.3 177	6	CQ207184	CQ207184 Sequence
400	35	79.5 214189	8	BS000022	BS000022 Pan trogl	c 473	34	77.3 177	6	CQ228652	CQ228652 Sequence
401	35	79.5 214258	14	AC132994	AC132994 Rattus no	c 474	34	77.3 177	6	CQ230641	CQ230641 Sequence
c 402	35	79.5 215757	9	AC105780	AC105780 Rattus no	c 475	34	77.3 177	6	CQ238545	CQ238545 Sequence
c 403	35	79.5 215758	9	AC122398	AC122398 Mus muscu	c 476	34	77.3 177	6	CQ266808	CQ266808 Sequence
c 404	35	79.5 216506	14	AC107127	AC107127 Rattus no	c 477	34	77.3 177	6	CQ268777	CQ268777 Sequence
c 405	35	79.5 217334	9	AC123695	AC123695 Mus muscu	c 478	34	77.3 177	6	CQ276181	CQ276181 Sequence
c 406	35	79.5 222521	14	AC121383	AC121383 Rattus no	c 479	34	77.3 177	6	CQ303762	CQ303762 Sequence
c 407	35	79.5 222879	14	AC105727	AC105727 Rattus no	c 480	34	77.3 177	6	CQ305800	CQ305800 Sequence
c 408	35	79.5 224817	14	AC111476	AC111476 Rattus no	c 481	34	77.3 177	6	CQ313175	CQ313175 Sequence
c 409	35	79.5 226356	14	AC106919	AC106919 Rattus no	c 482	34	77.3 177	6	CQ341092	CQ341092 Sequence
410	35	79.5 228030	14	AC164232	AC164232 Bos tauru	c 483	34	77.3 177	6	CQ342973	CQ342973 Sequence
411	35	79.5 228414	14	AC128563	AC128563 Rattus no	c 484	34	77.3 186	6	CQ081904	CQ081904 Sequence
412	35	79.5 229428	14	AC094493	AC094493 Rattus no	c 485	34	77.3 186	6	CQ155292	CQ155292 Sequence
c 413	35	79.5 230215	14	AC152328	AC152328 Bos tauru	c 486	34	77.3 186	6	CQ187708	CQ187708 Sequence
c 414	35	79.5 231166	14	AC134074	AC134074 Rattus no	c 487	34	77.3 186	6	CQ238499	CQ238499 Sequence
c 415	35	79.5 232127	14	AC113724	AC113724 Rattus no	c 488	34	77.3 186	6	CQ313118	CQ313118 Sequence
416	35	79.5 232726	14	AC159787	AC159787 Bos tauru	c 489	34	77.3 186	6	CQ350655	CQ350655 Sequence
c 417	35	79.5 233707	14	AC157070	AC157070 Bos tauru	c 490	34	77.3 227	5	AB125730	AB125730 Cyclomys
c 418	35	79.5 234379	14	AC098908	AC098908 Rattus no	c 491	34	77.3 251	10	BV193014	BV193014 sqmml7644
c 419	35	79.5 234655	14	AC162235	AC162235 Bos tauru	c 492	34	77.3 351	1	YPS270406	AJ270406 Yersinia
c 420	35	79.5 235607	14	AC102597	AC102597 Mus muscu	c 493	34	77.3 351	1	YPS270407	AJ270407 Yersinia
c 421	35	79.5 235812	14	AC094700	AC094700 Rattus no	c 494	34	77.3 351	1	YPS270408	AJ270408 Yersinia
c 422	35	79.5 235966	14	AC096147	AC096147 Rattus no	c 495	34	77.3 390	15	RICSINE05	DB5049 Oryza sativ
c 423	35	79.5 238463	14	AC120066	AC120066 Rattus no	c 496	34	77.3 408	6	CQ053416	CQ053416 Sequence
c 424	35	79.5 238896	14	AC137226	AC137226 Rattus no	c 497	34	77.3 408	6	CQ068555	CQ068555 Sequence
425	35	79.5 238980	14	AC106296	AC106296 Rattus no	c 498	34	77.3 408	6	CQ095622	CQ095622 Sequence
c 426	35	79.5 239746	14	AC125557	AC125557 Rattus no	c 499	34	77.3 408	6	CQ134366	CQ134366 Sequence
c 427	35	79.5 240209	14	AC130920	AC130920 Rattus no	c 500	34	77.3 408	6	CQ172868	CQ172868 Sequence
c 428	35	79.5 240498	9	AB121692	AB121692 Mus muscu	c 501	34	77.3 408	6	CQ202062	CQ202062 Sequence
c 429	35	79.5 240992	14	AC0931467	AC0931467 Rattus no	c 502	34	77.3 408	6	CQ217630	CQ217630 Sequence
c 430	35	79.5 242056	14	AC096280	AC096280 Rattus no	c 503	34	77.3 408	6	CQ256191	CQ256191 Sequence
c 431	35	79.5 242163	14	AC106264	AC106264 Rattus no	c 504	34	77.3 408	6	CQ293278	CQ293278 Sequence
c 432	35	79.5 243163	14	AC117093	AC117093 Rattus no	c 505	34	77.3 408	6	CQ330234	CQ330234 Sequence
c 433	35	79.5 245531	6	CQ870385	CQ870385 Sequence	c 506	34	77.3 457	6	BD119575	BD119575 EST and e
434	35	79.5 249522	14	AC115404	AC115404 Rattus no	c 507	34	77.3 457	6	AR424022	AR424022 Sequence
c 435	35	79.5 250298	14	AC093975	AC093975 Rattus no	c 508	34	77.3 457	6	AX984716	AX984716 Sequence
c 436	35	79.5 254114	14	AC129431	AC129431 Rattus no	c 509	34	77.3 475	6	CQ051576	CQ051576 Sequence
c 437	35	79.5 254372	14	AC132548	AC132548 Rattus no	c 510	34	77.3 475	6	CQ066628	CQ066628 Sequence
438	35	79.5 254518	14	AC160161	AC160161 Bos tauru	c 511	34	77.3 475	6	CQ093677	CQ093677 Sequence
c 439	35	79.5 254839	14	AC094263	AC094263 Rattus no	c 512	34	77.3 475	6	CQ132451	CQ132451 Sequence
c 440	35	79.5 255176	14	AC094295	AC094295 Rattus no	c 513	34	77.3 475	6	CQ171021	CQ171021 Sequence
c 441	35	79.5 255270	14	AC130767	AC130767 Rattus no	c 514	34	77.3 475	6	CQ200165	CQ200165 Sequence

c 515	34	77.3	475	6	CQ215650	Sequence	588	34	77.3	2323	15	AY150803	Arabidops
c 516	34	77.3	475	6	CQ254244	Sequence	589	34	77.3	2374	15	AK119748	Oryza sat
c 517	34	77.3	475	6	CQ291257	Sequence	590	34	77.3	2394	15	AF223573	Nicotiana
c 518	34	77.3	475	6	CQ28374	Sequence	591	34	77.3	2430	15	AB114834	Physcomit
c 519	34	77.3	531	10	BV312332	Sequence	c 592	34	77.3	2596	8	BC043429	BC043429
c 520	34	77.3	569	6	CQ072968	Sequence	593	34	77.3	2712	15	AY394079	AY394079
c 521	34	77.3	569	6	CQ142598	Sequence	594	34	77.3	2794	8	HSM808659	HSM808659
c 522	34	77.3	569	6	CQ178084	Sequence	595	34	77.3	2916	15	AK072141	AK072141
c 523	34	77.3	569	6	CQ225833	Sequence	596	34	77.3	2965	15	AF280748	AF280748
c 524	34	77.3	569	6	CQ300956	Sequence	597	34	77.3	3637	2	BT003535	BT003535
c 525	34	77.3	569	6	CQ338241	Sequence	598	34	77.3	3891	15	AY692025	AY692025
c 526	34	77.3	577	6	CQ103862	Sequence	599	34	77.3	4807	6	CQ582287	CQ582287
c 527	34	77.3	577	6	CQ225905	Sequence	c 600	34	77.3	5000	6	AR526835	AR526835
c 528	34	77.3	577	6	CQ263959	Sequence	601	34	77.3	5213	15	AY394078	AY394078
c 529	34	77.3	577	6	CQ301044	Sequence	602	34	77.3	6365	6	AR304138	AR304138
c 530	34	77.3	583	10	BV377877	Sequence	603	34	77.3	6443	15	AY059631	AY059631
c 531	34	77.3	687	10	BV069987	Sequence	604	34	77.3	6672	8	AF132360	AF132360
c 532	34	77.3	689	15	AY218824	Sequence	c 605	34	77.3	6765	15	AY166604	AY166604
c 533	34	77.3	704	10	BV613714	Sequence	606	34	77.3	6954	8	HSM806693	HSM806693
c 534	34	77.3	717	10	BV558656	Sequence	c 607	34	77.3	8757	9	MUSAL1PA	MUSAL1PA
c 535	34	77.3	720	8	HS32281A	Sequence	c 608	34	77.3	10065	8	AF240785	AF240785
c 536	34	77.3	725	10	BV640704	Sequence	609	34	77.3	10214	1	AE011477	AE011477
c 537	34	77.3	736	8	HSB334161	Sequence	610	34	77.3	10433	1	AE005168	AE005168
c 538	34	77.3	738	10	BV558831	Sequence	c 611	34	77.3	10797	1	AE004339	AE004339
c 539	34	77.3	760	6	AR494810	Sequence	c 612	34	77.3	10902	1	AE011858	AE011858
c 540	34	77.3	760	6	AX281095	Sequence	613	34	77.3	10976	13	AF098735	AF098735
c 541	34	77.3	772	10	BV606597	Sequence	614	34	77.3	11726	1	AE013807	AE013807
c 542	34	77.3	778	10	BV562130	Sequence	615	34	77.3	12237	15	AC007658	AC007658
c 543	34	77.3	798	6	AX2111369	Sequence	c 616	34	77.3	12788	8	AP000290	AP000290
c 544	34	77.3	870	6	AX434204	Sequence	c 617	34	77.3	19557	13	AY029400	AY029400
c 545	34	77.3	891	6	AX472318	Sequence	c 618	34	77.3	21006	8	AY855291	AY855291
c 546	34	77.3	1013	2	AF079446	Sequence	c 619	34	77.3	21335	8	AB050248	AB050248
c 547	34	77.3	1127	6	AX958407	Sequence	c 620	34	77.3	23447	8	AC125230	AC125230
c 548	34	77.3	1243	15	JRE312231	Sequence	c 621	34	77.3	24578	8	AY299456	AY299456
c 549	34	77.3	1358	8	HSM802039	Sequence	622	34	77.3	24985	1	AV372755	AV372755
c 550	34	77.3	1435	15	AB031202	Sequence	623	34	77.3	25621	1	PSY54826	PSY54826
c 551	34	77.3	1746	6	AX815466	Sequence	c 624	34	77.3	26591	1	AE008794	AE008794
c 552	34	77.3	1746	15	AY040054	Sequence	625	34	77.3	28735	2	L16622	L16622
c 553	34	77.3	1746	15	AF108123	Sequence	c 626	34	77.3	29605	15	AC074395	AC074395
c 554	34	77.3	1776	15	AY053422	Sequence	627	34	77.3	32958	14	AC110295	AC110295
c 555	34	77.3	1794	15	BT008358	Sequence	c 628	34	77.3	35058	6	CQ577817	CQ577817
c 556	34	77.3	1894	15	GMU41473	Sequence	c 629	34	77.3	40270	14	AC164594	AC164594
c 557	34	77.3	1940	15	STPLC2	Sequence	c 630	34	77.3	43034	6	AX695905	AX695905
c 558	34	77.3	1946	15	AF434168	Sequence	631	34	77.3	40893	3	AC150249	AC150249
c 559	34	77.3	1978	6	AR278871	Sequence	632	34	77.3	42149	15	SPAC1B3	SPAC1B3
c 560	34	77.3	1980	6	AR1164	Sequence	633	34	77.3	42198	14	AC166542	AC166542
c 561	34	77.3	1980	6	BD075118	Sequence	634	34	77.3	43238	8	AL954328	AL954328
c 562	34	77.3	1980	6	BD075120	Sequence	635	34	77.3	43441	8	HSJ9613	HSJ9613
c 563	34	77.3	2006	15	AY093217	Sequence	636	34	77.3	45695	14	AC087610	AC087610
c 564	34	77.3	2006	15	VU085250	Sequence	637	34	77.3	57991	15	AC018721	AC018721
c 565	34	77.3	2009	15	STPLC3	Sequence	c 638	34	77.3	61618	9	AY143165	AY143165
c 566	34	77.3	2032	15	AF360206	Sequence	c 639	34	77.3	65142	14	AC121501	AC121501
c 567	34	77.3	2038	15	DSA291467	Sequence	640	34	77.3	67970	14	AC117714	AC117714
c 568	34	77.3	2069	6	CQ850809	Sequence	c 641	34	77.3	67970	14	AC117714	AC117714
c 569	34	77.3	2069	8	AK127991	Sequence	642	34	77.3	70820	14	AC017451	AC017451
c 570	34	77.3	2079	15	GMU25027	Sequence	643	34	77.3	72013	8	AL356974	AL356974
c 571	34	77.3	2098	15	AF223351	Sequence	644	34	77.3	73368	8	AL391239	AL391239
c 572	34	77.3	2114	15	FSPLC	Sequence	645	34	77.3	75382	14	AC023242	AC023242
c 573	34	77.3	2134	15	GMU41474	Sequence	646	34	77.3	75430	5	CR626868	CR626868
c 574	34	77.3	2139	6	AR494816	Sequence	647	34	77.3	76447	6	AX116354	AX116354
c 575	34	77.3	2139	6	AX281101	Sequence	648	34	77.3	76798	6	AX111305	AX111305
c 576	34	77.3	2164	15	ATHATPLC2	Sequence	649	34	77.3	76798	8	HS3222B1	HS3222B1
c 577	34	77.3	2170	15	GMU41475	Sequence	c 650	34	77.3	76887	14	AC007775	AC007775
c 578	34	77.3	2175	6	CQ582288	Sequence	651	34	77.3	78261	8	AC117463	AC117463
c 579	34	77.3	2183	15	DQ113467	Sequence	c 652	34	77.3	81672	15	AB020755	AB020755
c 580	34	77.3	2195	15	DQ113456	Sequence	c 653	34	77.3	81735	5	CR847782	CR847782
c 581	34	77.3	2195	15	DQ113457	Sequence	c 654	34	77.3	81904	8	HS276A23	HS276A23
c 582	34	77.3	2195	15	DQ113458	Sequence	c 655	34	77.3	82584	15	NCB1383	NCB1383
c 583	34	77.3	2207	15	AF332874	Sequence	c 656	34	77.3	85574	6	CQ861540	CQ861540
c 584	34	77.3	2218	15	AK064924	Sequence	c 657	34	77.3	86574	8	HS83387	HS83387
c 585	34	77.3	2276	6	AR494815	Sequence	c 658	34	77.3	87854	8	AL360014	AL360014
c 586	34	77.3	2276	6	AX281100	Sequence	c 659	34	77.3	88361	14	AC138077	AC138077
c 587	34	77.3	2307	15	AK070452	Sequence	c 660	34	77.3	89171	8	AC008404	AC008404

661	34	77.3	89582	5	CR388169	CR388169 Zebrafish	734	34	77.3	119077	8	AL354982	AL354982 Human DNA
662	34	77.3	89747	8	AL591022	AL591022 Human DNA	c 735	34	77.3	119667	14	AC145417	AC145417 Felis cat
663	34	77.3	90858	14	AP008185	AP008185 Lotur cor	736	34	77.3	122168	8	AC127383	AC127383 Homo sapi
664	34	77.3	92946	14	AC140020_3	Continuation (4 of	737	34	77.3	123013	14	AC010005	AC010005 Drosophil
665	34	77.3	95111	15	ATF27K19_9	Continuation (4 of	738	34	77.3	124283	14	AC108368	AC108368 Pan trogl
666	34	77.3	95356	8	AL139112	AL139112 Human DNA	c 739	34	77.3	126503	8	AC008114	AC008114 Homo sapi
667	34	77.3	95448	8	AC005572	AC005572 Homo sapi	c 740	34	77.3	127426	15	AC137075	AC137075 Genomic s
668	34	77.3	98441	15	AP007292	AP007292 Lotur cor	741	34	77.3	127824	8	AL592205	AL592205 Human DNA
669	34	77.3	100000	8	AP000042	AP000042 Homo sapi	742	34	77.3	129619	15	CNS07YOA	AL731763 Oryza sat
670	34	77.3	100000	8	AP000110	AP000110 Homo sapi	743	34	77.3	129688	8	HS197B24	AL49217 Homo sapi
671	34	77.3	100000	8	AP000186	AP000186 Homo sapi	744	34	77.3	130824	8	AC078880	AC078880 Homo sapi
672	34	77.3	100886	14	AP003831	AP003831 Oryza sat	745	34	77.3	131241	14	AC114893	AC114893 Felis cat
673	34	77.3	103319	8	AC004217	AC004217 Homo sapi	746	34	77.3	132933	14	AC152154	AC152154 Dasyypus n
674	34	77.3	103762	8	AL354891	AL354891 Human DNA	c 747	34	77.3	133673	4	AC146891	AC146891 Dasyypus n
675	34	77.3	104762	14	AC108086	AC108086 Homo sapi	c 748	34	77.3	134928	8	AC091394	AC091394 Homo sapi
676	34	77.3	105133	8	AC010130	AC010130 Homo sapi	c 749	34	77.3	137213	8	AC005343	AC005343 Homo sapi
677	34	77.3	105823	14	AC019871	AC019871 Drosophil	c 750	34	77.3	137519	9	AL731658	AL731658 Mouse DNA
678	34	77.3	105948	9	BX511139	BX511139 Mouse DNA	751	34	77.3	137570	15	AC136999	AC136999 Oryza sat
679	34	77.3	107098	15	AP006660	AP006660 Lotur cor	752	34	77.3	138209	8	HS0J4421	AL078597 Human DNA
680	34	77.3	108032	8	AL162421	AL162421 Human DNA	c 753	34	77.3	139966	8	AC004820	AC004820 Homo sapi
681	34	77.3	108711	8	AC097523	AC097523 Homo sapi	754	34	77.3	140041	15	CR382279	CR382279 M truncat
682	34	77.3	110000	1	AB000516_17	Continuation (18 o	755	34	77.3	140462	15	AC145219	AC145219 Medicago
683	34	77.3	110000	1	AB006470_09	Continuation (10 o	c 756	34	77.3	141815	8	AC112716	AC112716 Homo sapi
684	34	77.3	110000	1	CR543861_30	Continuation (31 o	c 757	34	77.3	141930	8	AC146389	AC146389 Pan trogl
685	34	77.3	110000	1	CR626927_17	Continuation (18 o	758	34	77.3	141944	8	HS0J90G1	AL132670 Human DNA
686	34	77.3	110000	1	CR628336_15	Continuation (16 o	759	34	77.3	142335	14	AC136164	AC136164 Rattus no
687	34	77.3	110000	1	AB017220_22	Continuation (23 o	760	34	77.3	142924	14	CR762428	CR762428 Danio rer
688	34	77.3	110000	1	AB017282_03	Continuation (4 of	c 761	34	77.3	143515	15	AP003252	AP003252 Oryza sat
689	34	77.3	110000	1	AB017333_01	Continuation (2 of	c 762	34	77.3	143790	8	AC055811	AC055811 Homo sapi
690	34	77.3	110000	1	AB017354_15	Continuation (16 o	763	34	77.3	144455	15	AC129720	AC129720 Oryza sat
691	34	77.3	110000	1	AB017354_16	Continuation (17 o	764	34	77.3	144659	8	AC006145	AC006145 Homo sapi
692	34	77.3	110000	1	AB017354_32	Continuation (33 o	765	34	77.3	145246	9	AC115906	AC115906 Mus muscu
693	34	77.3	110000	1	AP008716_13	Continuation (14 o	c 766	34	77.3	145578	9	AC126548	AC126548 Mus muscu
694	34	77.3	110000	1	AP008934_11	Continuation (12 o	c 767	34	77.3	145992	8	AC134919	AC134919 Homo sapi
695	34	77.3	110000	1	BA000004_20	Continuation (21 o	c 768	34	77.3	146174	14	AC138823	AC138823 Homo sapi
696	34	77.3	110000	1	BA000026_01	Continuation (2 of	769	34	77.3	146245	8	AC148388	AC148388 Rhinoloph
697	34	77.3	110000	1	BX936398_24	Continuation (25 o	770	34	77.3	148667	8	HS1018E9	AL035455 Human DNA
698	34	77.3	110000	1	BX936398_25	Continuation (26 o	c 771	34	77.3	149015	14	AC022437	AC022437 Homo sapi
699	34	77.3	110000	1	CP000002_01	Continuation (9 of	c 772	34	77.3	149022	14	AC135162	AC135162 Medicago
700	34	77.3	110000	1	CP000026_08	Continuation (9 of	c 773	34	77.3	149317	8	AL954214	AL954214 Pan trogl
701	34	77.3	110000	1	CP000058_45	Continuation (46 o	774	34	77.3	149700	14	AP002738	AP002738 Homo sapi
702	34	77.3	110000	1	CP000075_19	Continuation (20 o	775	34	77.3	150311	9	AC112081	AC112081 Rattus no
703	34	77.3	110000	1	CP000075_33	Continuation (34 o	776	34	77.3	150315	14	BX324220	BX324220 Danio rer
704	34	77.3	110000	1	CP000089_12	Continuation (13 o	777	34	77.3	150755	9	AC154274	AC154274 Mus muscu
705	34	77.3	110000	6	BD430793_08	Continuation (9 of	778	34	77.3	150812	8	AC004765	AC004765 Homo sapi
706	34	77.3	110000	14	AC098746_0	AC098746 Rattus no	c 779	34	77.3	150812	9	BX649549	BX649549 Mouse DNA
707	34	77.3	110000	14	AC106253_2	Continuation (3 of	c 780	34	77.3	151087	9	AC144589	AC144589 Homo sapi
708	34	77.3	110000	14	AC110828_1	Continuation (2 of	c 781	34	77.3	151242	14	AC144589	AC144589 Homo sapi
709	34	77.3	110000	14	AC139803_1	Continuation (2 of	c 782	34	77.3	152244	8	AP000352	AP000352 Homo sapi
710	34	77.3	110000	14	AC140148_2	Continuation (3 of	783	34	77.3	152952	8	AC074143	AC074143 Homo sapi
711	34	77.3	110000	14	AC156266_05	Continuation (3 of	c 784	34	77.3	153125	14	AC148432	AC148432 Sorex ara
712	34	77.3	110000	14	BX813324_1	Continuation (2 of	c 785	34	77.3	153187	8	AL954215	AL954215 Pan trogl
713	34	77.3	110000	14	CT005272_18	Continuation (13 o	786	34	77.3	153607	14	AC024615	AC024615 Homo sapi
714	34	77.3	110000	15	AP008216_136	Continuation (19 o	787	34	77.3	153907	8	AC146761	AC146761 Pan trogl
715	34	77.3	110000	15	AP008216_137	Continuation (138	c 788	34	77.3	154084	15	AP003734	AP003734 Oryza sat
716	34	77.3	110000	15	AP008216_208	Continuation (138	789	34	77.3	154288	8	AL513164	AL513164 Human DNA
717	34	77.3	110000	15	AP008218_005	Continuation (209	790	34	77.3	154854	14	AC152387	AC152387 Dasyypus n
718	34	77.3	110000	15	AP008218_174	Continuation (6 of	791	34	77.3	155164	8	AC005165	AC005165 Homo sapi
719	34	77.3	110000	15	CR382139_03	Continuation (175	c 792	34	77.3	155236	14	AC009701	AC009701 Homo sapi
720	34	77.3	110000	15	AB016817_11	Continuation (4 of	793	34	77.3	155358	5	BX323839	BX323839 Zebrafish
721	34	77.3	110000	15	AB017350_04	Continuation (12 o	c 794	34	77.3	155450	8	AC005951	AC005951 Homo sapi
722	34	77.3	110000	15	AP008207_354	Continuation (5 of	c 795	34	77.3	155663	14	AC152857	AC152857 Ateleiix
723	34	77.3	110000	15	AP008209_099	Continuation (355	796	34	77.3	155666	8	AC008696	AC008696 Homo sapi
724	34	77.3	110000	15	AP008211_014	Continuation (100	797	34	77.3	156441	8	AC091869	AC091869 Homo sapi
725	34	77.3	110000	15	AP008211_169	Continuation (15 o	798	34	77.3	156836	14	AC159960	AC159960 Rhinoloph
726	34	77.3	110000	15	AP008213_294	Continuation (170	c 799	34	77.3	157119	5	BX276185	BX276185 Zebrafish
727	34	77.3	110073	9	AL928567	Continuation (295	c 800	34	77.3	158091	8	AC140171	AC140171 Homo sapi
728	34	77.3	112353	8	AC109928	AL928567 Mouse DNA	c 801	34	77.3	158566	8	AC093694	AC093694 Homo sapi
729	34	77.3	115859	15	CNS08C80	AC109928 Homo sapi	802	34	77.3	158566	14	AC137915	AC137915 Felis cat
730	34	77.3	116814	8	AR526865	BX000510 Oryza sat	803	34	77.3	158921	8	AC024329	AC024329 Homo sapi
731	34	77.3	118999	6	AR526865	AR526865 Homo sapi	804	34	77.3	159347	14	AC138971	AC138971 Homo sapi
732	34	77.3	118999	8	AP240786	AR526865 Sequence	805	34	77.3	159888	9	AC107698	AC107698 Mus muscu
733	34	77.3	118999	8	AP000351	AP240786 Homo sapi	806	34	77.3	160551	2	AC008320	AC008320 Drosophil

807	34	77.3	161004	14	AC163401	AC163401 Mus muscu
c 808	34	77.3	161649	14	AC023652	AC023652 Homo sapi
c 809	34	77.3	161810	9	AL663086	AL663086 Mouse DNA
c 810	34	77.3	161827	8	AC103816	AC103816 Homo sapi
c 811	34	77.3	161841	8	AC135279	AC135279 Homo sapi
c 812	34	77.3	162194	14	AC026592	AC026592 Homo sapi
c 813	34	77.3	162195	14	AC009906	AC009906 Homo sapi
c 814	34	77.3	162328	8	AP000350	AP000350 Homo sapi
c 815	34	77.3	162473	5	AL929518	AL929518 Zebrafish
c 816	34	77.3	162662	8	AC011927	AC011927 Homo sapi
c 817	34	77.3	162734	9	AC109200	AC109200 Mus muscu
c 818	34	77.3	162931	8	AC083862	AC083862 Homo sapi
c 819	34	77.3	163186	8	AB045360	AB045360 Homo sapi
c 820	34	77.3	163388	15	AP005199	AP005199 Oryza sat
c 821	34	77.3	163811	14	AC108689	AC108689 Homo sapi
c 822	34	77.3	165346	14	AC141422	AC141422 Pan trogl
c 823	34	77.3	165346	14	AC141422	AC141422 Pan trogl
c 824	34	77.3	165743	5	CR759748	CR759748 Zebrafish
c 825	34	77.3	166653	8	AC129851	AC129851 Homo sapi
c 826	34	77.3	166869	8	AC016395	AC016395 Homo sapi
c 827	34	77.3	167198	14	AC148509	AC148509 Macropus
c 828	34	77.3	167441	14	AC073832	AC073832 Homo sapi
c 829	34	77.3	167464	14	AC108563	AC108563 Papio anu
c 830	34	77.3	168114	8	AL354828	AL354828 Human DNA
c 831	34	77.3	168140	14	AC008248	AC008248 Homo sapi
c 832	34	77.3	169199	8	AC090497	AC090497 Homo sapi
c 833	34	77.3	169280	14	AC161860	AC161860 Ateles
c 834	34	77.3	170560	13	AF689437	AF689437 Deerpox v
c 835	34	77.3	170706	8	AC113389	AC113389 Homo sapi
c 836	34	77.3	172276	8	AC139795	AC139795 Homo sapi
c 837	34	77.3	172599	5	BX890576	BX890576 Zebrafish
c 838	34	77.3	172637	6	AX686783	AX686783 Sequence
c 839	34	77.3	172742	15	AC027659	AC027659 Oryza sat
c 840	34	77.3	173028	8	AC090371	AC090371 Homo sapi
c 841	34	77.3	173387	8	AC025041	AC025041 Homo sapi
c 842	34	77.3	173670	8	AL133351	AL133351 Human DNA
c 843	34	77.3	174128	14	AP001123	AP001123 Homo sapi
c 844	34	77.3	174176	5	AL772340	AL772340 Zebrafish
c 845	34	77.3	174453	14	AC105041	AC105041 Bos tauru
c 846	34	77.3	174640	14	AC024479	AC024479 Homo sapi
c 847	34	77.3	175092	9	AC122491	AC122491 Mus muscu
c 848	34	77.3	175115	2	AC105899	AC105899 Drosophila
c 849	34	77.3	175302	8	AL355677	AL355677 Human DNA
c 850	34	77.3	175819	9	AC102290	AC102290 Mus muscu
c 851	34	77.3	176140	14	AC118959	AC118959 Rattus no
c 852	34	77.3	176206	14	AC130425	AC130425 Homo sapi
c 853	34	77.3	176641	14	AC084869	AC084869 Homo sapi
c 854	34	77.3	177250	9	AC087782	AC087782 Mus muscu
c 855	34	77.3	177364	9	AC156274	AC156274 Mus muscu
c 856	34	77.3	178022	15	AC025783	AC025783 Oryza sat
c 857	34	77.3	178560	14	AC134877	AC134877 Homo sapi
c 858	34	77.3	179536	14	AC010814	AC010814 Homo sapi
c 859	34	77.3	179548	9	AC110200	AC110200 Mus muscu
c 860	34	77.3	179668	9	AC124399	AC124399 Mus muscu
c 861	34	77.3	180136	1	BAC180K	BAC180K B. subtilis
c 862	34	77.3	180346	8	AC092615	AC092615 Homo sapi
c 863	34	77.3	180953	14	AC069269	AC069269 Homo sapi
c 864	34	77.3	180982	8	AC078985	AC078985 Homo sapi
c 865	34	77.3	180999	8	AL513205	AL513205 Human DNA
c 866	34	77.3	181569	14	AC148577	AC148577 Rhinoph
c 867	34	77.3	182162	14	AC025302	AC025302 Homo sapi
c 868	34	77.3	182339	14	AC092761	AC092761 Papio anu
c 869	34	77.3	182715	8	AC138819	AC138819 Homo sapi
c 870	34	77.3	182898	2	AC010113	AC010113 Drosophila
c 871	34	77.3	183219	14	AC134443	AC134443 Mus muscu
c 872	34	77.3	183222	14	BX511220	BX511220 Danio rer
c 873	34	77.3	183551	9	AC134570	AC134570 Mus muscu
c 874	34	77.3	183991	9	AC122340	AC122340 Mus muscu
c 875	34	77.3	184340	14	AC155201	AC155201 Callithri
c 876	34	77.3	184984	14	AC112056	AC112056 Rattus no
c 877	34	77.3	185162	8	AC140132	AC140132 Homo sapi
c 878	34	77.3	185432	14	AC141200	AC141200 Rattus no
c 879	34	77.3	185796	14	AC138857	AC138857 Homo sapi
880	34	77.3	185866	9	AC109510	AC109510 Mus muscu
c 881	34	77.3	185918	2	AC009839	AC009839 Drosophila
c 882	34	77.3	185936	14	AC139837	AC139837 Homo sapi
c 883	34	77.3	185986	14	AC140123	AC140123 Homo sapi
c 884	34	77.3	186239	8	AC112642	AC112642 Homo sapi
c 885	34	77.3	187600	14	AC092515	AC092515 Papio anu
c 886	34	77.3	187628	8	AC011311	AC011311 Homo sapi
c 887	34	77.3	188259	14	AC161591	AC161591 Mus muscu
c 888	34	77.3	188356	14	AC009550	AC009550 Homo sapi
c 889	34	77.3	188977	14	AC140146	AC140146 Homo sapi
c 890	34	77.3	189076	14	AC023788	AC023788 Homo sapi
c 891	34	77.3	189627	14	AC128926	AC128926 Rattus no
c 892	34	77.3	189984	5	AL844883	AL844883 Zebrafish
c 893	34	77.3	190462	8	AP003057	AP003057 Homo sapi
c 894	34	77.3	190874	14	AC139830	AC139830 Homo sapi
c 895	34	77.3	191142	14	AC015467	AC015467 Homo sapi
c 896	34	77.3	191490	14	AC023656	AC023656 Homo sapi
c 897	34	77.3	192163	9	AC125359	AC125359 Mus muscu
c 898	34	77.3	192226	14	AC109000	AC109000 Rattus no
c 899	34	77.3	192885	14	AC139910	AC139910 Rattus no
c 900	34	77.3	192902	14	AC007480	AC007480 Homo sapi
c 901	34	77.3	192993	14	AC145108	AC145108 Homo sapi
c 902	34	77.3	193981	14	AC027542	AC027542 Homo sapi
c 903	34	77.3	194388	14	AC118853	AC118853 Rattus no
c 904	34	77.3	194763	14	AC150852	AC150852 Bos tauru
c 905	34	77.3	195205	14	AC131416	AC131416 Rattus no
c 906	34	77.3	196484	14	AC129435	AC129435 Rattus no
c 907	34	77.3	196555	14	AC155825	AC155825 Mus muscu
c 908	34	77.3	197866	14	AC138895	AC138895 Homo sapi
c 909	34	77.3	197972	8	AC138865	AC138865 Homo sapi
c 910	34	77.3	197980	14	AC166872	AC166872 Oryctolag
c 911	34	77.3	198230	8	AC092862	AC092862 Homo sapi
c 912	34	77.3	198768	15	AY678298	AY678298 Lycoperes
c 913	34	77.3	199414	5	BX323990	BX323990 Zebrafish
c 914	34	77.3	199799	14	AC141576	AC141576 Rattus no
c 915	34	77.3	199808	9	AC103600	AC103600 Mus muscu
c 916	34	77.3	200167	8	AC002480	AC002480 Homo sapi
c 917	34	77.3	200883	14	AL590107	AL590107 Homo sapi
c 918	34	77.3	201259	9	AC159504	AC159504 Mus muscu
c 919	34	77.3	201262	14	AC118512	AC118512 Rattus no
c 920	34	77.3	201638	14	AC144987	AC144987 Homo sapi
c 921	34	77.3	201791	14	CR774178	CR774178 Danio rer
c 922	34	77.3	202559	14	AC153031	AC153031 Bos tauru
c 923	34	77.3	202576	9	AC140931	AC140931 Mus muscu
c 924	34	77.3	202726	5	AL773593	AL773593 Zebrafish
c 925	34	77.3	202751	14	BX248237	BX248237 Danio rer
c 926	34	77.3	202766	14	AC141507	AC141507 Rattus no
c 927	34	77.3	203681	14	AC122079	AC122079 Rattus no
c 928	34	77.3	203901	1	BSUB0021	Z99124 Bacillus su
c 929	34	77.3	204249	14	CT010462	CT010462 Mus muscu
c 930	34	77.3	204707	14	AC124732	AC124732 Mus muscu
c 931	34	77.3	204891	9	AC158307	AC158307 Mus muscu
c 932	34	77.3	205360	14	AC159055	AC159055 Bos tauru
c 933	34	77.3	205476	5	CR753899	CR753899 Zebrafish
c 934	34	77.3	206024	9	AC153547	AC153547 Mus muscu
c 935	34	77.3	206196	14	AC135295	AC135295 Rattus no
c 936	34	77.3	207534	14	AC024698	AC024698 Homo sapi
c 937	34	77.3	207694	14	AC117086	AC117086 Rattus no
c 938	34	77.3	208332	9	AC122227	AC122227 Mus muscu
c 939	34	77.3	209920	14	AC163665	AC163665 Mus muscu
c 940	34	77.3	210518	14	AC128317	AC128317 Rattus no
c 941	34	77.3	211542	14	AC073206	AC073206 Homo sapi
c 942	34	77.3	211595	14	AC019334	AC019334 Homo sapi
c 943	34	77.3	211668	14	AC115146	AC115146 Rattus no
c 944	34	77.3	212042	14	AC130593	AC130593 Rattus no
c 945	34	77.3	212399	8	AC084784	AC084784 Homo sapi
c 946	34	77.3	212426	8	AC104335	AC104335 Homo sapi
c 947	34	77.3	212590	9	BX547995	BX547995 Mouse DNA
c 948	34	77.3	213230	9	AC159472	AC159472 Mus muscu
c 949	34	77.3	214405	9	AC121871	AC121871 Mus muscu
c 950	34	77.3	214864	14	AC162627	AC162627 Bos tauru
c 951	34	77.3	215494	14	AC161831	AC161831 Bos tauru
c 952	34	77.3	215962	14	AC112905	AC112905 Rattus no

```
c 953 34 77.3 216311 9 AC156982
c 954 34 77.3 216422 14 AC096528
c 955 34 77.3 216546 14 AC138863
c 956 34 77.3 216844 14 AC130788
c 957 34 77.3 216877 8 AC106791
c 958 34 77.3 217068 9 AC126459
c 959 34 77.3 217636 14 AC116260
c 960 34 77.3 217950 14 AC146652
c 961 34 77.3 217993 14 AC109864
c 962 34 77.3 218135 14 AC099250
c 963 34 77.3 218675 14 AC015847
c 964 34 77.3 218893 14 AC150547
c 965 34 77.3 219221 14 AC122278
c 966 34 77.3 219723 14 AC152289
c 967 34 77.3 220669 14 AC123017
c 968 34 77.3 220720 14 AC154400
c 969 34 77.3 220926 14 AC064858
c 970 34 77.3 221011 14 AC158305
c 971 34 77.3 221074 8 AC090503
c 972 34 77.3 221402 14 AC015995
c 973 34 77.3 221567 14 AC141219
c 974 34 77.3 221617 5 BX530018
c 975 34 77.3 222536 14 AC117919
c 976 34 77.3 223099 5 BX548071
c 977 34 77.3 224037 9 AL670399
c 978 34 77.3 225381 14 AC162417
c 979 34 77.3 225655 9 AC099690
c 980 34 77.3 227488 14 AC133316
c 981 34 77.3 227529 14 AC145052
c 982 34 77.3 228273 14 AC111408
c 983 34 77.3 228310 14 AC106556
c 984 34 77.3 228327 14 CR456625
c 985 34 77.3 230360 14 AC164196
c 986 34 77.3 230635 14 AC128319
c 987 34 77.3 230705 14 AC130996
c 988 34 77.3 231213 14 AC111561
c 989 34 77.3 231406 14 AC132626
c 990 34 77.3 231675 9 AL591410
c 991 34 77.3 231741 14 AC110855
c 992 34 77.3 231914 14 CR752646
c 993 34 77.3 232477 14 AC152433
c 994 34 77.3 232495 14 AC106594
c 995 34 77.3 233763 14 AC119297
c 996 34 77.3 234237 9 AC138272
c 997 34 77.3 234515 14 AC098639
c 998 34 77.3 234560 14 AC097419
c 999 34 77.3 234714 9 AC110176
c1000 34 77.3 234889 14 AC098605
```

ALIGNMENTS

```
RESULT 1
AX467373 AX467373 1260 bp DNA linear PAT 16-JUL-2002
LOCUS Sequence 3 from Patent WO0238612.
DEFINITION AX467373
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
SOURCE Felis sp.
ORGANISM Felis sp.
REFERENCE 1
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1260
/organism="Felis sp."
/mol_type="unassigned DNA"
```

```
/db_xref="taxon:9687"
ORIGIN
Alignment Scores: 5.21 Length: 1260
Pred. No.: 44.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 6 Gaps: 0
DB:
US-10-774-176-5 (1-9) x AX467373 (1-1260)
Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 286 TTCTCACCAGCAATCAGTGGCGGTG 312
RESULT 2
AX821533 1260 bp DNA linear PAT 10-DEC-2003
LOCUS Sequence 1 from Patent WO03068816.
DEFINITION AX821533
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS Felis catus (cat)
SOURCE Felis catus
ORGANISM Felis catus
REFERENCE 1
AUTHORS Carroll,M.M., Kingman,S.M. and Redchenko,I.M.
TITLE MHC class I peptide epitopes from the human 5t4 tumor-associated antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores: 5.21 Length: 1260
Pred. No.: 44.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 6 Gaps: 0
DB:
US-10-774-176-5 (1-9) x AX821533 (1-1260)
Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 286 TTCTCACCAGCAATCAGTGGCGGTG 312
RESULT 3
AX821548 1260 bp DNA linear PAT 10-DEC-2003
LOCUS Sequence 1 from Patent WO03068815.
DEFINITION AX821548
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS Felis catus (cat)
SOURCE Felis catus
ORGANISM Felis catus
REFERENCE 1
AUTHORS Carroll,M.O., Harrop,R.O. and Kingman,S.O.
TITLE MHC class II peptide epitope of 5t4 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
```

```

FEATURES
  source
    Location/Qualifiers
      1..1260
        /organism="Felis catus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9685"
ORIGIN
Alignment Scores:
  Pred. No.: 5.21 Length: 1260
  Score: 44.00 Matches: 9
  Percent Similarity: 100.0% Conservative: 0
  Best Local Similarity: 100.0% Mismatches: 0
  Query Match: 100.0% Indels: 0
  DB: 6 Gaps: 0
US-10-774-176-5 (1-9) x AX821548 (1-1260)
Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
Db 286 TTCCTCACCGGCATCAGCTGCCGTG 312
RESULT 4
BD249731
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1263)
  Carroll,M.W. and Myers,K.A.
  Polypeptide
  Patent: WO 0029428-A 1 25-MAY-2000;
  CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
  BIOMEDICA LTD (GB)
  Location/Qualifiers
    1..1263
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
  Pred. No.: 5.23 Length: 1263
  Score: 44.00 Matches: 9
  Percent Similarity: 100.0% Conservative: 0
  Best Local Similarity: 100.0% Mismatches: 0
  Query Match: 100.0% Indels: 0
  DB: 6 Gaps: 0
US-10-774-176-5 (1-9) x AX025011 (1-1263)
Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
Db 289 TTCCTTACCGGCACACCGCTGCCGTG 315
RESULT 6
AX149553
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE
  1 Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
  Ellard,F.M. and Myers,K.A.
  Antibodies
  Patent: WO 0136486-A 14 25-MAY-2001;
  Oxford Biomedica (UK) Limited (GB)
  Location/Qualifiers
    1..1263
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="5T4"
ORIGIN
Alignment Scores:
  Pred. No.: 5.23 Length: 1263
  Score: 44.00 Matches: 9
  Percent Similarity: 100.0% Conservative: 0
  Best Local Similarity: 100.0% Mismatches: 0
  Query Match: 100.0% Indels: 0
  DB: 6 Gaps: 0
US-10-774-176-5 (1-9) x AX149553 (1-1263)
Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
Db 289 TTCCTTACCGGCACACCGCTGCCGTG 315
FEATURES
  source
    Location/Qualifiers
      1..1260
        /organism="Felis catus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9685"
ORIGIN
Alignment Scores:
  Pred. No.: 5.21 Length: 1260
  Score: 44.00 Matches: 9
  Percent Similarity: 100.0% Conservative: 0
  Best Local Similarity: 100.0% Mismatches: 0
  Query Match: 100.0% Indels: 0
  DB: 6 Gaps: 0
US-10-774-176-5 (1-9) x AX821548 (1-1260)
Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
Db 286 TTCCTCACCGGCATCAGCTGCCGTG 312
RESULT 4
BD249731
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1263)
  Carroll,M.W. and Myers,K.A.
  Polypeptide
  Patent: JP 2002530060-A 1 17-SEP-2002;
  OXFORD BIOMEDICA LTD
  OS Homo sapiens (human)
  PN JP 2002530060-A/1
  PD 17-SEP-2002
  PF 18-NOV-1999 JP 2000582415
  PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
  30-JUL-1999 GB 9917995.4
  PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
  PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K14/065,
  C07K19/00,
  PC C12N15/00,
  CC Polypeptide
  FH Key
  FT source
    Location/Qualifiers
      1..1263
        /organism="Homo sapiens (human)"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
  Pred. No.: 5.23 Length: 1263
  Score: 44.00 Matches: 9
  Percent Similarity: 100.0% Conservative: 0
  Best Local Similarity: 100.0% Mismatches: 0
  Query Match: 100.0% Indels: 0
  DB: 6 Gaps: 0
US-10-774-176-5 (1-9) x BD249731 (1-1263)
Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
Db 289 TTCCTTACCGGCACACCGCTGCCGTG 315

```


Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 Db 289 TTCTTACGGGCAACCACTGGCGGTG 315

RESULT 7
 LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
 DEFINITION Sequence 1 from Patent EP1160323.
 ACCESSION AX316086
 VERSION AX316086.1 GI:17899278
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Carroll,M.W. and Myers,K.A.
 TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
 JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1. .1263
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 5.23 Length: 1263
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x AX316086 (1-1263)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 Db 289 TTCTTACGGGCAACCACTGGCGGTG 315

RESULT 8
 LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
 DEFINITION Sequence 1 from Patent WO0238612.
 ACCESSION AX467371
 VERSION AX467371.1 GI:21900602
 KEYWORDS
 SOURCE Canis sp.
 ORGANISM Canis sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.

REFERENCE
 AUTHORS Myers,K., Drury,N. and Carroll,M.
 TITLE Polypeptide
 JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1. .1263
 /organism="Canis sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:9616"

ORIGIN

Alignment Scores:
 Pred. No.: 5.23 Length: 1263
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x AX467371 (1-1263)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 Db 289 TTCTTACGGGCAACCACTGGCGGTG 315

RESULT 9
 LOCUS CQ731678 2053 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 17612 from Patent WO02068579.
 ACCESSION CQ731678
 VERSION CQ731678.1 GI:42308932
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 17612 06-SEP-2002;
 PE Corporation (NY) (US)

FEATURES
 source
 1. .2053
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 8.34 Length: 2053
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x CQ731678 (1-2053)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
 Db 373 TTCTTACGGGCAACCACTGGCGGTG 399

RESULT 10
 LOCUS HS5T4OA 2053 bp RNA linear PRI 18-APR-2005
 DEFINITION Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.
 ACCESSION Z29083
 VERSION Z29083.1 GI:435654
 KEYWORDS 5T4 gene; 5T4 oncofoetal antigen.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Myers,K.A., Rahi-Saund,V., Davison,M.D., Young,J.A., Cheater,A.J.
 TITLE Isolation of a cDNA encoding 5T4 oncofoetal trophoblast
 glycoprotein. An antigen associated with metastasis contains
 leucine-rich repeats
 JOURNAL J. Biol. Chem. 269 (12), 9319-9324 (1994)

ORIGIN

Alignment Scores:
 Pred. No.: 8.34 Length: 2053
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

```

source
1. .2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62. .372
/product="LRR N-terminal flank"
/label=N-flank
85. .1347
/codon_start=1
/evidence=experimental
/product="SF4 oncofetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSSS
APFLASVSAQPLPDQCPALCECESEARTVKVNRNLTEVPTDLPAYVRLFLTGNO
LAVLPAGAFARPPLAEALALNLSGRSLDEVRAGAFELPSLRQLDLSHNPLADLSPF
AFSGSNASVSAPSLVELILNHIVPPEDERQNRSEFGVMVAALLAGRALQGLRLLELA
SNHFLYLPRLVLAQLPSLRHLDLNNSLSVSLTVSFRLNLTLESJLHLDNALKVLHNG
TLAEIQADLPHRIVFLDNNPWVCDCHMADMTWLKETEVVQGRDRLTCAYPERQNRVL
LELNSADLDQDLPILPSPQTSYVFLGILVALIGLPLVLYLNRKGKKMMINIRDAC
RDHMEGYHYRVEINADPRLTNLSNSDV"
130. .171
sig_peptide
misc_RNA
373. .966
/product="Leucine rich repeat region"
/label=LRRs
966. .1119
/misc_RNA
/product="LRR C-terminal flank"
/label=C-flank
1153. .1215
/misc_RNA
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

ORIGIN
Alignment Scores:
Pred. No.: 8.34 Length: 2053
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x HS5T40A (1-2053)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 373 TTCCTTACCGGACACGCTGCCGCTG 399

RESULT 11
BD127282
LOCUS BD127282 2359 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127282
VERSION BD127282.1 GI:23222227
KEYWORDS JP 2002017375-A/2713.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2359)
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.

TITLE BD127282
JOURNAL
FEATURES
source
1. .2359
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
424. .1575
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"
/translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSSS
APFLASVSAQPLPDQCPALCECESEARTVKVNRNLTEVPTDLPAYVRLFLTGNO
LAVLPAGAFARPPLAEALALNLSGRSLDEVRAGAFELPSLRQLDLSHNPLADLSPF
AFSGSNASVSAPSLVELILNHIVPPEDERQNRSEFGVMVAALLAGRALQGLRLLELA

CDS
1. .2359
/codon_start=1
/protein_id="CAF85958.1"
/db_xref="GI:45502668"
/translation="MPGCGSRGPAAGDGRRLRLARLALVLLGWSSSPTSSASSPSSS
APFLASVSAQPLPDQCPALCECESEARTVKVNRNLTEVPTDLPAYVRLFLTGNO
LAVLPAGAFARPPLAEALALNLSGRSLDEVRAGAFELPSLRQLDLSHNPLADLSPF
AFSGSNASVSAPSLVELILNHIVPPEDERQNRSEFGVMVAALLAGRALQGLRLLELA

```

SNHFLYLPDRVLAQLPSLRHLDLSNNLSVLSVYVPSRNLTHLESJLHLEDNALKVLHNG
TLAEQLGRLRVRFLDNPWVCDKMDMTWTKAEVVGCKRLTCAYPEKMRNVL
LELSADLDCDPLPPSLQTSYVFLGIVLALIGALFELVLYLNKGIKK"

ORIGIN

Alignment Scores:
Pred. No.: 9.53 Length: 2359
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x CQ782724 (1-2359)

Qy 1 PheLeuThrGlyAaGlnLeuAlaVal 9
Db 712 TTCCTTACCGCAACAGCTGGCGGTG 738

RESULT 13

AK074786 2359 bp mRNA linear PRI 03-SEP-2002
LOCUS Homo sapiens cDNA FLJ90305 fis, clone NT2RP2000694, highly similar
DEFINITION to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AK074786
VERSION AK074786.1 GI:22760460
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE

1 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T., Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S., Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Kojima, S., Nagehara, K., Masuko, Y., Ono, T., Okano, K., Yoshikawa, Y., Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and Ninomiya, K.
NEDO human cDNA sequencing project

TITLE

Unpublished

REFERENCE

2 (bases 1 to 2359)

AUTHORS

Isogai, T. and Otsuki, T.

TITLE

Direct Submission

JOURNAL

Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).

FEATURES

source

1..2359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="Cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 9.53 Length: 2359
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x AK074786 (1-2359)

Qy 1 PheLeuThrGlyAaGlnLeuAlaVal 9
Db 712 TTCCTTACCGCAACAGCTGGCGGTG 738

RESULT 14

BD127283 2361 bp DNA linear PAT 18-SEP-2002
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION
ACCESSION BD127283
VERSION BD127283.1 GI:23222228
KEYWORDS JP 2002017375-A/2714.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2361)
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL
COMMENT
OS Helix Research Institute
OS Homo sapiens (human)
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO KOGA, YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers
(426) ..(1685).

FEATURES

source

1..2361
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 9.53 Length: 2361
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x BD127283 (1-2361)

Qy 1 PheLeuThrGlyAaGlnLeuAlaVal 9
Db 714 TTCCTTACCGCAACAGCTGGCGGTG 740

RESULT 15

CQ782726 2361 bp DNA linear PAT 17-MAR-2004
LOCUS Sequence 2866 from Patent EP1396543.
DEFINITION
ACCESSION CQ782726
VERSION CQ782726.1 GI:45502669
KEYWORDS
SOURCE Homo sapiens (human)

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
REFERENCE
AUTHORS Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2866 10-MAR-2004;
RESEARCH Association for Biotechnology (JP)
FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
426..1688
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF85961.1"
/db_xref="GI:45502670"
/translation="MPGCGSRGPAAGDGRRLRLRLALVLLGWSSSSPTSSASSPSSS
APFLASAVSAQPLPDQCPALCESEAAARTVKVNRNLTEVPTDLPAVYRNPLFLTGNO
LAVLPAGAFARPPLAEIALNLGSRDLDEVRAGAFELHPSLRQLDLSHNPLADLSPP
AFSGSNASVAPSPLVELILNHIVPEDERQNRSPGCMVVAALLAGRLQGLRLLELA
SNHFLYPRDVLQAPSLRHLDSNLSVLTIVSFRNLTHLSLHSDNALKVLHNG
TLAEQLGPHIRVFLDNNPWCDCHMADMTWLKETEYVQGRDLTCAYPEKMRNVL
LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLVLYLNRRGIKKWMHNRDAD
RDHMEGYHYRYEINADPRLTNLSSSDV"

CDS
US-10-774-176-5 (1-9) x CQ782726 (1-2361)

Alignment Scores:
Pred. No.: 9.53 Length: 2361
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x Q782726 (1-2361)

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 714 TTCCTTACCGGCAACGAGTGGCCGTG 740

RESULT 16
LOCUS AX961916 2361 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 127 from Patent WO03104277.
ACCESSION AX961916
VERSION AX961916.1 GI:40881326
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
REFERENCE
AUTHORS Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 127 18-DEC-2003;
RESEARCH Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
426..1688
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06467.1"
/db_xref="GI:40881327"
/translation="MPGCGSRGPAAGDGRRLRLRLALVLLGWSSSSPTSSASSPSSS
APFLASAVSAQPLPDQCPALCESEAAARTVKVNRNLTEVPTDLPAVYRNPLFLTGNO
LAVLPAGAFARPPLAEIALNLGSRDLDEVRAGAFELHPSLRQLDLSHNPLADLSPP
AFSGSNASVAPSPLVELILNHIVPEDERQNRSPGCMVVAALLAGRLQGLRLLELA
SNHFLYPRDVLQAPSLRHLDSNLSVLTIVSFRNLTHLSLHSDNALKVLHNG
TLAEQLGPHIRVFLDNNPWCDCHMADMTWLKETEYVQGRDLTCAYPEKMRNVL
LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLVLYLNRRGIKKWMHNRDAD
RDHMEGYHYRYEINADPRLTNLSSSDV"

CDS
US-10-774-176-5 (1-9) x AX961916 (1-2361)

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 714 TTCCTTACCGGCAACGAGTGGCCGTG 740

RESULT 17
LOCUS AK074790 2361 bp mRNA linear PRI 09-JUL-2005
DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens ST4 oncofetal trophoblast glycoprotein gene.
ACCESSION AK074790
VERSION AK074790.1 GI:22760466
KEYWORDS Homo sapiens
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
REFERENCE
AUTHORS Otsuki, T., Ota, T., Nishikawa, T., Hayashi, K., Suzuki, Y.,
Yamamoto, J., Wakamatsu, A., Kimura, K., Sakamoto, K., Hatano, N.,
Kawai, Y., Ishii, S., Saito, K., Kojima, S., Sugiyama, T., Ono, T.,
Okano, K., Yoshikawa, Y., Aotaka, S., Sasaki, N., Hattori, A.,
Okumura, K., Nagai, K., Sugano, S. and Isogai, T.
TITLE Signal Sequence and Keyword Trap in silico for Selection of
Full-Length Human cDNAs Encoding Secretion or Membrane Proteins
from Oligo-Capped cDNA Libraries
JOURNAL DNA Res. 12, 117-126 (2005)
REFERENCE
AUTHORS Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahari, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotaka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2361)
AUTHORS Isogai, T. and Otsuki, T.
DIRECT Submission
JOURNAL Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection; Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1..2361
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

APFLASAVSAQPLPDQCPALCESEAAARTVKVNRNLTEVPTDLPAVYRNPLFLTGNO
LAVLPAGAFARPPLAEIALNLGSRDLDEVRAGAFELHPSLRQLDLSHNPLADLSPP
AFSGSNASVAPSPLVELILNHIVPEDERQNRSPGCMVVAALLAGRLQGLRLLELA
SNHFLYPRDVLQAPSLRHLDSNLSVLTIVSFRNLTHLSLHSDNALKVLHNG
TLAEQLGPHIRVFLDNNPWCDCHMADMTWLKETEYVQGRDLTCAYPEKMRNVL
LELNSADLDCDPIPLPSLQTSYVFLGIVLALIGAILFLVLYLNRRGIKKWMHNRDAD
RDHMEGYHYRYEINADPRLTNLSSSDV"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 9.53 Length: 2361
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

```

US-10-774-176-5 (1-9) x AX961916 (1-2361)

```

QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db 714 TTCCTTACCGGCAACGAGTGGCCGTG 740

```

RESULT 17
AK074790

LOCUS AK074790 2361 bp mRNA linear PRI 09-JUL-2005
DEFINITION Homo sapiens cDNA FLJ90309 fis, clone NT2RP2000903, highly similar
to Homo sapiens ST4 oncofetal trophoblast glycoprotein gene.

ACCESSION AK074790

VERSION AK074790.1 GI:22760466

KEYWORDS Homo sapiens

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE DNA Res. 12, 117-126 (2005)

AUTHORS

TITLE

JOURNAL NEDO human cDNA sequencing project

REFERENCE Unpublished

AUTHORS 3 (bases 1 to 2361)

TITLE

JOURNAL Direct Submission

COMMENT

Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba, 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection; Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

source

1..2361

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

```

/clones="NT2RP2000903"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/notes="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN

Alignment Scores:
Pred. No.: 9.53 Length: 2361
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x AK074790 (1-2361)

QY 1 PheLeuThrGlyAenGlnLeuAlaVal 9
Db 714 TTCTTACCGCAACACGAGTGGCGGTG 741

RESULT 18
LOCUS BC037161 2379 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens trophoblast glycoprotein, mRNA (cDNA clone MGC:15317
IMAGE:4138906), complete cds.
ACCESSION BC037161
VERSION BC037161.2 GI:33872201
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2379)
Straussberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schweizer, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.B., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
PUBMED 2 (bases 1 to 2379)
Straussberg, R.
Direct Submission
TITLE Submitted (03-SEP-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT On Aug 19, 2003 this sequence version replaced gi:22713382.
Contact: MGC help desk
Email: gcgape-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <http://www.nisc.nih.gov/> Contact: nisc.mcngri.nih.gov Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Series: IRAL Plate: 26 Row: m Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5729717.

FEATURES

source Location/Qualifiers

1..2379

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:15317 IMAGE:4138906"

/tissue_type="Muscle, rhabdomyosarcoma"

/clone_lib="NIH_MGC_17"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1..2379

/gene="TPBG"

/note="synonyms: M6P1, 5T4-AG, 5T4"

/db_xref="GeneID:7162"

427..1689

/gene="TPBG"

/codon_start=1

/product="5T4 oncofetal trophoblast glycoprotein"

/protein_id="AAH37161.1"

/db_xref="GI:22713383"

/db_xref="GeneID:7162"

/db_xref="MIM:190920"

/translation="MPGCSRGAAGDGLRLRLALVLLGWSSSPTSSASFSFSS APFLASVASQPPPLDQPCALCESEARTVCYNRLTEVPTDLPAYVNLFLTGNQ LAVPAGAFARRPPLAEALNLSGRSLDEVRAGAFELPRLQDLGNPLADLSPP AFSGNSASVSPSLVELLNHIYPPEQRNSRFGVMVAALLAGRALQGLRLLELA SNHFLYLPRLVLAQLPSRLHLDNNSLSVLTYSFRNLTHLESILHLDNALKVLHNG TLAEQLGLPHIRVFLDNNPWCDCMADMTLKTETVVGQKRLTCAYPEKRRVL LELNSADLDCDPLPPSLQTSYVFLGVLALIGALFLVLVLYLNRRGKIKKWHNIRDAC RDHMEGHYRIEINADPRLTNLSNDV"

ORIGIN

Alignment Scores:

Pred. No.: 9.6 Length: 2379

Score: 44.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x BC037161 (1-2379)

QY 1 PheLeuThrGlyAenGlnLeuAlaVal 9

Db 715 TTCTTACCGCAACACGAGTGGCGGTG 741

RESULT 19

AB168308

LOCUS

DEFINITION Macaca fascicularis testis cDNA clone: QtsA-11109, similar to human trophoblast glycoprotein (TPBG), mRNA, RefSeq: NM_006670.3.

ACCESSION AB168308

VERSION AB168308.1 GI:67967899


```
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA09930.1"
/db_xref="GI:3805947"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGSRGPAAGDGRRLRLRLALVILGWSSSPTSSASSPSS
APFLASVASQPLPDQCPALCESEAAATVKVNRNLTEVTDLPAYVNRFLITGNQ
LAVLPAGAFARPPFLAEALNLGSRLEDEVRAGAFELPSLRQLDLSHNPDLADLSP
AFSGSNASVASPSPLEILNHIIVPDEQRQRNFEQGVVAALAGRALQGLRLLELA
SNFLYLPRDVLAQLPSLRHLDSNLSVLTYSFRNLTHLSLHLEDNALKVLHNG
TLAEQLGPHIRVFLDNNPWCDCHMADMTLAKETVVOGKORLTCAYPEKMRNVL
LELSADLDCDPIPLPSPSQTSYVFLGIVLALIGALFLLVLYNRKGIKKWNHNRDAC
RDHMEGHYHYREINADPRLTNLSNSDV"
3431..3516
/gene="574"
3517..4690
/gene="574"
/product="5T4 oncofetal trophoblast glycoprotein"
/gene="574"
5331..5336
/polyA_signal
5380..5385
/polyA_signal
/gene="574"

ORIGIN
Alignment Scores:
Pred. No.: 21.7 Length: 5551
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-5 (1-9) x HSA012159 (1-5551)

Qy 1 PhleuThrGlyAenGlnLeuAlaVal 9
Db 3719 TTCTTACCGCAACCAAGCTGCGCGTG 3745

RESULT 21
HSJ492P14 121909 bp DNA linear PRI 18-MAY-2005
LOCUS Human DNA sequence from clone RP3-492P14 on chromosome 6q13-15
DEFINITION Contains a single stranded DNA binding protein pseudogene, the TPBG
gene for trophoblast glycoprotein (5T4-AG) and a CpG island, the
complete sequence.
AUI21977
VERSION AUI21977.11 GI:11863678
KEYWORDS HTG; CpG island; TPBG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 121909)
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Dec 15, 2000 this sequence version replaced gi:11558491.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emn, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C/elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
```

```
RP3-492P14 is from the library RPCI-3 constructed by the group of
Pier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1..121909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="q13-15"
/clone="RP3-492P14"
/clone_lib="RPCI-3"
misc_feature
100
/note="Clone right end: RP1-93K22"
complement(10004..10982)
/locus_tag="RP3-492P14.2-001"
/pseudo
CDS
complement(10004..10982)
/locus_tag="RP3-492P14.2-001"
/note="match: proteins: P81877 Q99LX9 Q9BWM6 Q9CYZ8 Q9D6L4
Q9P038 Q9Y4T7"
/pseudo
misc_feature
86539
/note="Clone left end: RP1-90G1"
109639..116836
/gene="TPBG"
/mRNA
join(109639..109916,110631..116836)
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
/product="trophoblast glycoprotein"
/note="match: ESTs: AA149121 AA152323 AA565852 AA643734
AL544610 AW471072 AW662538 BE260089 BF306457 BF306926
BF314984 BI196133 BI562387 BM069633 BM670613
match: CDNAs: AJ420536.1 Z29083.1"
110970..112232
/gene="TPBG"
/locus_tag="RP3-492P14.1-001"
/standard_name="OTTHUMP00000016786"
/note="match: proteins: Q13641 Q9QYD9 Q9ZOL0"
/codon_start=1
/product="trophoblast glycoprotein"
/protein_id="CAI21546.1"
/db_xref="GI:56203539"
/db_xref="Gene:12004"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="UniProt/TREMBL:Q13641"
/translation="MPGCGSRGPAAGDGRRLRLRLALVILGWSSSPTSSASSPSS
APFLASVASQPLPDQCPALCESEAAATVKVNRNLTEVTDLPAYVNRFLITGNQ
LAVLPAGAFARPPFLAEALNLGSRLEDEVRAGAFELPSLRQLDLSHNPDLADLSP
AFSGSNASVASPSPLEILNHIIVPDEQRQRNFEQGVVAALAGRALQGLRLLELA
SNFLYLPRDVLAQLPSLRHLDSNLSVLTYSFRNLTHLSLHLEDNALKVLHNG
TLAEQLGPHIRVFLDNNPWCDCHMADMTLAKETVVOGKORLTCAYPEKMRNVL
LELSADLDCDPIPLPSPSQTSYVFLGIVLALIGALFLLVLYNRKGIKKWNHNRDAC
```

RDHMEGYHYRYEINADPRLTNLSSNSDV"
 116817 -116822
 /gene="TPBG"
 /locus_tag="RP3-492P14.1-001"
 116836
 /gene="TPBG"
 /locus_tag="RP3-492P14.1-001"
 121909
 /note="Clone_right_end: RP3-492P14"

ORIGIN

Alignment Scores:
 Pred. No.: 422 Length: 121909
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-5. (1-9) x HSJ492P14 (1-121909)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

Db 111258 TTCTTACCGGCACACAGCTGGCCGTG 111284

RESULT 22

CR854913_3

WPCOMMENT

Sequence split into 4 fragments LOCUS CR854913 Accession CR854913

Fragment Name Begin End

CR854913_0 1 110000

CR854913_1 100001 210000

CR854913_2 200001 310000

CR854913_3 300001 388345

Continuation (4 of 4) of CR854913 from base 300001 (CR854913 Danio rerio clone DKXY-18P1

Alignment Scores:

Pred. No.: 1.52e+03 Length: 88345
 Score: 41.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 93.2% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-5 (1-9) x CR854913_3 (1-88345)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

Db 77142 TTCTTACCGGTAATCAGCTCTCGGTA 77168

RESULT 23

CR854985

LOCUS CR854985 168825 bp DNA linear HTG 13-MAY-2005

DEFINITION Danio rerio clone DKXY-42017, *** SEQUENCING IN PROGRESS ***, 12

unordered pieces.

ACCESSION CR854985

VERSION CR854985.3 GI:64972214

KEYWORDS HTG; HTGS; PHASE1.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 168825)

REFERENCE Sime,S.

Direct Submission

Submitted (12-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On May 13, 2005 this sequence version replaced gi:55058488.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>
 Contact: zfish-help@sanger.ac.uk
 ----- Project Information
 Center project name: zk42017
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 164928 bases at least Q40
 Consensus quality: 165688 bases at least Q30
 Consensus quality: 166375 bases at least Q20
 Insert size: 167725; sum-of-contigs
 Insert size: 177628; 3.4% error; agarose-fp
 Quality coverage: 6.25x in Q20 bases; sum-of-contigs Quality
 coverage: 6.02x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1
 * 20921: contig of 20921 bp in length
 * 20922
 * 21021: gap of 100 bp
 * 21022
 * 28602: contig of 7581 bp in length
 * 28603
 * 28702: gap of 100 bp
 * 28703
 * 37137: contig of 8435 bp in length
 * 37138
 * 37237: gap of 100 bp
 * 37238
 * 68400: contig of 31163 bp in length
 * 68401
 * 68500: gap of 100 bp
 * 68501
 * 89611: contig of 21111 bp in length
 * 89612
 * 89711: gap of 100 bp
 * 89712
 * 101523: contig of 11812 bp in length
 * 101524
 * 101623: gap of 100 bp
 * 101624
 * 107502: contig of 5879 bp in length
 * 107503
 * 107602: gap of 100 bp
 * 107603
 * 128891: contig of 21289 bp in length
 * 128892
 * 128991: gap of 100 bp
 * 128992
 * 143731: contig of 14740 bp in length
 * 143732
 * 143831: gap of 100 bp
 * 143832
 * 151795: contig of 7964 bp in length
 * 151796
 * 151895: gap of 100 bp
 * 151896
 * 165144: contig of 13249 bp in length
 * 165145
 * 165244: gap of 100 bp
 * 165245
 * 168825: contig of 3581 bp in length.

FEATURES

source

1. .168825
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /db_xref="taxon:7955"
 /clone="DKXY-42017"
 /clone_lib="DanioKey"

1. .20921

/note="assembly fragment:00742

fragment chain:1"

21022. .28602

/note="assembly fragment:00159

fragment chain:1"

28703. .37137

/note="assembly fragment:00343

fragment chain:1"

37238. .68400

/note="assembly fragment:01495

fragment chain:1"

68501. .89611

/note="assembly fragment:01223

fragment chain:1"

89712. .101523

/note="assembly fragment:00457

fragment chain:1"

101624. .107502

/note="assembly fragment:00102

fragment_chain:1"
misc_feature 107603..128891
/note="assembly_fragment:00966
fragment_chain:1"
misc_feature 128992..143731
/note="assembly_fragment:00586
fragment_chain:1"
misc_feature 143832..151795
/note="assembly_fragment:00056
fragment_chain:2"
misc_feature 151896..165144
/note="assembly_fragment:00243
fragment_chain:2"
misc_feature 165245..168825
/note="assembly_fragment:00025.0"

ORIGIN

Alignment Scores: 168825
Pred. No.: 2.84e+03 Length: 8
Score: 41.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.2% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-5 (1-9) x CR854985 (1-168825)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9
Db 13984 TTCCTACGGGTAATCAGCTCGGTA 14010

RESULT 24

AX829164 AX829164 927 bp DNA linear PAT 12-DEC-2003
LOCUS
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE 1
AUTHORS Mack,D.H.; Gish,K.C. and Afar,D.
TITLE Methods of diagnosis of breast cancer, compositions and methods of
screening for modulators of breast cancer

JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;

EOS Biotechnology, Inc. (US)

FEATURES
Location/Qualifiers

source 1..927
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 927
Pred. No.: 32.5 Length: 8
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.9% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-5 (1-9) x AX829164 (1-927)

Qy 1 PheLeuThrGlyAenGlnLeuAla 8

Db 289 TTCCTACCGGCAACCGCTGGCC 312

RESULT 25

AP008208.073

WPCOMMENT

Sequence split into 360 fragments LOCUS AP008208 Accession AP008208
Fragment Name Begin End
AP008208_000 1 110000
AP008208_001 100001 210000
AP008208_002 200001 310000
AP008208_003 300001 410000
AP008208_004 400001 510000
AP008208_005 500001 610000
AP008208_006 600001 710000
AP008208_007 700001 810000
AP008208_008 800001 910000
AP008208_009 900001 1010000
AP008208_010 1000001 1110000
AP008208_011 1100001 1210000
AP008208_012 1200001 1310000
AP008208_013 1300001 1410000
AP008208_014 1400001 1510000
AP008208_015 1500001 1610000
AP008208_016 1600001 1710000
AP008208_017 1700001 1810000
AP008208_018 1800001 1910000
AP008208_019 1900001 2010000
AP008208_020 2000001 2110000
AP008208_021 2100001 2210000
AP008208_022 2200001 2310000
AP008208_023 2300001 2410000
AP008208_024 2400001 2510000
AP008208_025 2500001 2610000
AP008208_026 2600001 2710000
AP008208_027 2700001 2810000
AP008208_028 2800001 2910000
AP008208_029 2900001 3010000
AP008208_030 3000001 3110000
AP008208_031 3100001 3210000
AP008208_032 3200001 3310000
AP008208_033 3300001 3410000
AP008208_034 3400001 3510000
AP008208_035 3500001 3610000
AP008208_036 3600001 3710000
AP008208_037 3700001 3810000
AP008208_038 3800001 3910000
AP008208_039 3900001 4010000
AP008208_040 4000001 4110000
AP008208_041 4100001 4210000
AP008208_042 4200001 4310000
AP008208_043 4300001 4410000
AP008208_044 4400001 4510000
AP008208_045 4500001 4610000
AP008208_046 4600001 4710000
AP008208_047 4700001 4810000
AP008208_048 4800001 4910000
AP008208_049 4900001 5010000
AP008208_050 5000001 5110000
AP008208_051 5100001 5210000
AP008208_052 5200001 5310000
AP008208_053 5300001 5410000
AP008208_054 5400001 5510000
AP008208_055 5500001 5610000
AP008208_056 5600001 5710000
AP008208_057 5700001 5810000
AP008208_058 5800001 5910000
AP008208_059 5900001 6010000
AP008208_060 6000001 6110000
AP008208_061 6100001 6210000
AP008208_062 6200001 6310000
AP008208_063 6300001 6410000
AP008208_064 6400001 6510000
AP008208_065 6500001 6610000
AP008208_066 6600001 6710000
AP008208_067 6700001 6810000
AP008208_068 6800001 6910000
AP008208_069 6900001 7010000
AP008208_070 7000001 7110000

AP008208_071 7210000
AP008208_072 7310000
AP008208_073 7300001
AP008208_074 7400001
AP008208_075 7500001
AP008208_076 7600001
AP008208_077 7700001
AP008208_078 7800001
AP008208_079 7900001
AP008208_080 8000001
AP008208_081 8100001
AP008208_082 8200001
AP008208_083 8300001
AP008208_084 8400001
AP008208_085 8500001
AP008208_086 8600001
AP008208_087 8700001
AP008208_088 8800001
AP008208_089 8900001
AP008208_090 9000001
AP008208_091 9100001
AP008208_092 9200001
AP008208_093 9300001
AP008208_094 9400001
AP008208_095 9500001
AP008208_096 9600001
AP008208_097 9700001
AP008208_098 9800001
AP008208_099 9900001
AP008208_100 10000001
AP008208_101 10100000
AP008208_102 10200001
AP008208_103 10300001
AP008208_104 10400001
AP008208_105 10500001
AP008208_106 10600001
AP008208_107 10700001
AP008208_108 10800001
AP008208_109 10900001
AP008208_110 11000001
AP008208_111 11100001
AP008208_112 11200001
AP008208_113 11300001
AP008208_114 11400001
AP008208_115 11500001
AP008208_116 11600001
AP008208_117 11700001
AP008208_118 11800001
AP008208_119 11900001
AP008208_120 12000001
AP008208_121 12100001
AP008208_122 12200001
AP008208_123 12300001
AP008208_124 12400001
AP008208_125 12500001
AP008208_126 12600001
AP008208_127 12700001
AP008208_128 12800001
AP008208_129 12900001
AP008208_130 13000001
AP008208_131 13100001
AP008208_132 13200001
AP008208_133 13300001
AP008208_134 13400001
AP008208_135 13500001
AP008208_136 13600001
AP008208_137 13700001
AP008208_138 13800001
AP008208_139 13900001
AP008208_140 14000001
AP008208_141 14100001
AP008208_142 14200001
AP008208_143 14300001

AP008208_144 14400001
AP008208_145 14500001
AP008208_146 14600001
AP008208_147 14700001
AP008208_148 14800001
AP008208_149 14900001
AP008208_150 15000001
AP008208_151 15100001
AP008208_152 15200001
AP008208_153 15300001
AP008208_154 15400001
AP008208_155 15500001
AP008208_156 15600001
AP008208_157 15700001
AP008208_158 15800001
AP008208_159 15900001
AP008208_160 16000001
AP008208_161 16100001
AP008208_162 16200001
AP008208_163 16300001
AP008208_164 16400001
AP008208_165 16500001
AP008208_166 16600001
AP008208_167 16700001
AP008208_168 16800001
AP008208_169 16900001
AP008208_170 17000001
AP008208_171 17100001
AP008208_172 17200001
AP008208_173 17300001
AP008208_174 17400001
AP008208_175 17500001
AP008208_176 17600001
AP008208_177 17700001
AP008208_178 17800001
AP008208_179 17900001
AP008208_180 18000001
AP008208_181 18100001
AP008208_182 18200001
AP008208_183 18300001
AP008208_184 18400001
AP008208_185 18500001
AP008208_186 18600001
AP008208_187 18700001
AP008208_188 18800001
AP008208_189 18900001
AP008208_190 19000001
AP008208_191 19100001
AP008208_192 19200001
AP008208_193 19300001
AP008208_194 19400001
AP008208_195 19500001
AP008208_196 19600001

Alignment Scores:

Pred. No.: 3.2e+03 Length: 110000
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.9% Indels: 0
DB: 15 Gaps: 0

US-10-774-176-5 (1-9) x AP008208_073 (1-110000)

QY 1 PheLeuThrGlyAsnGlnLeuAla 8
|||||
Db 37755 TTTTAACTGTTAACCCAGCTAGCT 37778

RESULT 26

AP005756

LOCUS

DEFINITION

AP005756 136267 bp DNA linear
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
BAC clone:OSJNB0035N08.

ACCESSION	AF005756	misc_feature	complement (6895..7272)
VERSION	AP005756.3 GI:49388934		/gene="OSJNBb0035N08.2"
KEYWORDS			/note="hypothetical ORF"
SOURCE	Oryza sativa (japonica cultivar-group)		predicted by GlimmerM
ORGANISM	Oryza sativa (japonica cultivar-group)	gene	this category is not included in IRGSP standard"
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		complement (11805..13042)
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	mRNA	/gene="OSJNBb0035N08.3"
	Ehrhartoideae; Oryzeae; Oryza.		complement (join(<11805..12148,12855..>13042))
REFERENCE	1		/gene="OSJNBb0035N08.3"
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.	misc_feature	/note="supported by full-length cDNA(s): AK111070"
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC		complement (11805..13042)
	Clone:OSJNBb0035N08		/gene="OSJNBb0035N08.3"
JOURNAL	Published Only in Database (2002)		/note="contains full-length cDNA(s): AK111070"
REFERENCE	2 (bases 1 to 136267)		non-coding transcript
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.	gene	probably inactive due to no initiation codon in CDS"
TITLE	Direct Submission		complement (13420..13668)
JOURNAL	Submitted (18-SEP-2002) Takuji Sasaki, National Institute of		/gene="OSJNBb0035N08.4"
	Agricultural Sciences, Rice Genome Research Program; Kamondai	misc_feature	complement (13420..13668)
	2-1-2, Teikuba, Ibaraki 305-8602, Japan		/gene="OSJNBb0035N08.4"
	(E-mail:tsaaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,	gene	/note="hypothetical ORF"
	Tel:81-298-7441, Fax:81-298-38-7468)		predicted by GlimmerM
COMMENT	On Jun 26, 2004 this sequence version replaced gi:42627749.		this category is not included in IRGSP standard"
	Genes were predicted from the integrated results of the following:	gene	complement (14428..14868)
	GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), PCENESH	mRNA	/gene="OSJNBb0035N08.5"
	(http://www.softberry.com/), GeneMark.hmm		complement (<14428..>14868)
	(http://opal.biology.gatech.edu/GeneMark/), GlimmerM	CDS	/gene="OSJNBb0035N08.5"
	(http://www.tigr.org/tdb/glimmer/gimr_form.html), RiceHMM		/note="start and end point are not identified"
	(http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor		complement (14428..14868)
	(http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4		/gene="OSJNBb0035N08.5"
	(http://globin.cse.psu.edu/html/docs/sim4.html), gap2		/note="predicted by GeneMark.hmm etc."
	(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The		/codon_start=1
	genomic sequence was searched against NCBI NonRedundant Protein		/product="hypothetical protein"
	database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA		/protein_id="BAD26155.1"
	sequence database at RGP or DBJ. Protein homologues of the coding		/db_xref="GI:49388935"
	regions were searched against NCBI NonRedundant Protein database		/translation="MPAVLLIVVVVVTVAAVLILVPLVVRGLAGLSAHSIHRH
	with BLASTP. ESTs represent the identified cDNA sequences using		RAGGASAAFLPHIRCHRSPFPDPPLSDPCGRWPPRLPTVGSATRLTIVGSV
	BLASTN with the corresponding DBJ accession no. and RGP clone ID.	gene	TAVASLRPLPLPSPSTAGGRRGRRSGGEEAAVVVALARR"
	Full-length cDNAs represent the identified cDNA sequences using		15873..16028
	BLASTN with the corresponding DBJ accession no.	misc_feature	/gene="OSJNBb0035N08.6"
	A gene with identity or significant homology to a protein is		15873..16028
	classified based on the protein name to indicate the homology level		/gene="OSJNBb0035N08.6"
	such as same name, 'putative-' and '-like protein'. A gene without		/note="hypothetical ORF"
	significant homology to any protein but with full-length cDNA or	gene	predicted by GlimmerM
	EST homology (covering almost the entire length of partial		this category is not included in IRGSP standard"
	sequence) is classified as an 'unknown' protein. A gene predicted		complement (13655..23497)
	by two or more gene prediction programs is classified as a	gene	/gene="OSJNBb0035N08.7"
	'hypothetical' protein according to IRGSP standard. A gene	mRNA	complement (join(19965..20252,20345..20405,20630..20777,
	predicted by a single gene prediction program is also classified as		21029..21282,22175..22283,22950..23497))
	a probable 'hypothetical' protein and is included as a		/gene="OSJNBb0035N08.7"
	miscellaneous feature of the sequence.		/note="supported by full-length cDNA(s): AK121870"
	The orientation of the sequence is from M13rev to -21M13 of the BAC	CDS	complement (join(20211..20252,20345..20405,20630..20777,
	clone. This sequence of OSJNBb0035N08 clone has an overlap with		21029..21282,22175..22283,22950..23238))
	P0620H05 (DBJ: AP005394) clone at 5' end and with QJ1711.D06		/gene="OSJNBb0035N08.7"
	(DBJ: AP004857) clone at 3' end. Detailed information on overlap		/note="contains EST(s): AU100805(C51513), C27285(C51513)
	and assembly quality together with annotation of this entry is		contains full-length cDNA(s): AK121870"
	available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.		/codon_start=1
FEATURES	Location/Qualifiers		/product="putative aux/IAA protein"
Source	1..136267		/protein_id="BAD26156.1"
	/organism="Oryza sativa (japonica cultivar-group)"		/db_xref="GI:49388936"
	/mol_type="genomic DNA"		/translation="MGEASESMKKISRGRLGSMGEPDSHRRHGDQEEBEKTLLELS
	/cultivar="Nipponbare"		LGLPQGWRAACRDKGTTKHSIAAAADDDGDKSSMLSIGYTLVSHSQKANKN
	/db_xref="taxon:39947"		KGSPREEAHPPTATGNALASNNNGCFQTSPTSPVVGWPPVPTFRNLATSSKASL
	/chromosome="2"		ELQDKKAABEIKRAPIKINMDGVPIGRKIDINAFDSTKLVSLADKCLFRGLLAA
	/clone="OSJNBb0035N08"		ORDPLTAGAKCCQEDVAISGLIDGTGYTLVYEDYEGDKVLGVDPWGMFVSVKRL
gene	complement (join(157..256,518..608,645..687,1665..1820))		RVLKTSLSLITSGRRKTAEC"
	/gene="OSJNBb0035N08.1"	gene	complement (24133..24630)
misc_feature	complement (join(157..256,518..608,645..687,1665..1820))		/gene="OSJNBb0035N08.8"
	/gene="OSJNBb0035N08.1"	misc_feature	complement (24133..24630)
	/note="hypothetical ORF"		/gene="OSJNBb0035N08.8"
	predicted by GENSCAN		/note="hypothetical ORF"
	this category is not included in IRGSP standard"		predicted by GlimmerM
gene	complement (6895..7272)		this category is not included in IRGSP standard"
	/gene="OSJNBb0035N08.2"		complement (27886..28131)

```

misc_feature
/gene="OSJNBb0035N08.9"
complement(27886..28131)
/gene="OSJNBb0035N08.9"
/note="hypothetical ORF
predicted by GlimmerM
this category is not included in IRGSP standard"
31945..32890
/gene="OSJNBb0035N08.10"
join(<31945..31975,32081..32221,32569..32627,
32831..>32890)
/gene="OSJNBb0035N08.10"
/note="start and end point are not identified"
join(31945..31975,32081..32221,32569..32627,32831..32890)
/gene="OSJNBb0035N08.10"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:49388937"
/translation="MACGLKNVGDRLPLCSGSTAVDORLCAEDRPPALSLAPASS
IASILTTRRYGGCKRDLRSYIAVPPKLPWAVVSITMBEQAGGDLVSL"
complement(33602..35288)
/gene="OSJNBb0035N08.11"
complement(join(33602..33842,33945..34031,34799..34969,
35069..35288))
/gene="OSJNBb0035N08.11"
/note="supported by full-length cDNA(s): AK058248"
complement(join(33810..33842,33945..34031,34799..34969,
35069..35194))
/gene="OSJNBb0035N08.11"
/note="contains EST(s): C99556(E20550)
contains full-length cDNA(s): AK058248"
/codon_start=1
/product="putative 40S ribosomal protein S24"
/protein_id="BAD26158.1"
/db_xref="GI:49388938"
/translation="MADTKTAPAVTLTRKFWNLLSRKQFVLEVLHPCGANYSKAD
LKEKLAKIYKVDKNCIFVFRTHFGGKSTGFGLIYDNLDAKKYEPKYLRLRNGL
ATKVKGRKQMKRNRKAKIRGVKTKAGDAGKKK"
35886..36086
/gene="OSJNBb0035N08.12"
<35886..36086
/gene="OSJNBb0035N08.12"
/note="start and end point are not identified"
35886..36086
/gene="OSJNBb0035N08.12"
/note="predicted by GENESH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD26159.1"
/db_xref="GI:49388939"

Alignment Scores:
Pred. No.: 3.93e+03 Length: 136267
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 90.9% Indels: 0
DB: 15 Gaps: 0

US-10-774-176-5 (1-9) x AP005756 (1-136267)

QY 1 PheLeuThrGlyAaGlnLeuAla 8
|||||
Db 113563 TTTTIACTGGTACCAGCTAGCT 113586

RESULT 27
AC083883
LOCUS AC083883 60199 bp DNA linear PRI 07-NOV-2001
DEFINITION Homo sapiens BAC clone RP11-792N18 from 7, complete sequence.
ACCESSION AC083883
VERSION AC083883.6 GI:14589790
KEYWORDS HTG.

```

```

SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
9847074
2 (bases 1 to 60199)
Goyea,E. and Abbott,A.
The sequence of Homo sapiens BAC clone RP11-792N18
Unpublished
3 (bases 1 to 60199)
Waterston,R.H.
Direct Submission
Submitted (04-OCT-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 60199)
Waterston,R.H.
Direct Submission
Submitted (03-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 60199)
Waterston,R.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 3, 2001 this sequence version replaced gi:14389375.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiense@wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0792N18

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)

```


ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
AUTHORS	1 (bases 1 to 209688)
TITLE	Submitted (21-APR-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	2 (bases 1 to 209688)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	On Jun 29, 2005 this sequence version replaced gi:62821922.
TITLE	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas
JOURNAL	assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
Center:	Genome Center
Center code:	BCM
Web site:	http://www.hgsc.bcm.tmc.edu/
Contact:	hgsc-help@bcm.tmc.edu
Project information	Center project name: PFEX
Center clone name:	CH240-79C10
Summary Statistics	Assembly program: Atlas 3.0;
Consensus quality:	202990 bases at least Q40
Consensus quality:	204529 bases at least Q30
Consensus quality:	205759 bases at least Q20
Estimated insert size:	207978; sum-of-contigs estimation
Quality coverage:	5x in Q20 bases; sum-of-contigs estimation
NOTE:	Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
NOTE:	This is a 'working draft' sequence. It currently consists of 16 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1	5278: contig of 5278 bp in length
5279	5328: gap of 50 bp
5329	10166: contig of 4838 bp in length
10167	10216: gap of 50 bp
10217	62997: contig of 52781 bp in length
63047	gap of 50 bp
63048	96337: contig of 33790 bp in length
96338	gap of unknown length
148433	contig of 51496 bp in length
148434	gap of 50 bp
171936	171935: contig of 23452 bp in length
171936	171985: gap of 50 bp
171986	176792: contig of 4807 bp in length
176793	176842: gap of 50 bp
186743	186761: contig of 9919 bp in length
186762	186811: gap of 50 bp
196723	196722: contig of 9911 bp in length
196723	197276: gap of 554 bp
198624	198623: contig of 1347 bp in length
198724	201270: gap of unknown length
201271	201585: gap of 315 bp
201586	203809: contig of 2224 bp in length
203810	203909: gap of unknown length
203910	205189: contig of 1280 bp in length
205190	205289: gap of unknown length
205290	206747: contig of 1458 bp in length
206748	206847: gap of unknown length
206848	208505: contig of 1658 bp in length
208506	208605: gap of unknown length
208606	209688: contig of 1083 bp in length.
Location/Qualifiers	1. .209688
organism="Bos taurus"	/mol_type="genomic DNA"
db_xref="taxon:9913"	/db_xref="taxon:9913"
clone="CH240-79C10"	/clone="CH240-79C10"
5279. .5328	/estimated_length=50
FEATURES	source
gap	gap

```

gap      10167. 10216
/estimated length=50
gap      62998. 63047
/estimated length=50
gap      96838. 96937
/estimated length=unknown
gap      148434. 148483
/estimated length=50
gap      171936. 171985
/estimated length=50
gap      176793. 176842
/estimated length=50
gap      186762. 186811
/estimated length=50
gap      196723. 197276
/estimated length=554
gap      198624. 198723
/estimated length=unknown
gap      201271. 201585
/estimated length=315
gap      203810. 203909
/estimated length=unknown
gap      205190. 205289
/estimated length=unknown
gap      206748. 206847
/estimated length=unknown
gap      208506. 208605
/estimated length=unknown

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.01e+04 Length: 209688
Score: 39.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 88.6% Indels: 0
DB: 14 Gaps: 0

```

US-10-774-176-5 (1-9) x AC150293 (1-209688)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9

Db 72429 TTCTTAACGGGAACCTCCAGCAGTC 72403

```

RESULT 29
AC126962
LOCUS
DEFINITION
AC126962 222692 bp DNA linear HTG 10-MAY-2003
Rattus norvegicus clone CH230-132K18, WORKING DRAFT SEQUENCE, 2
unordered pieces.

```

```

AC126962
VERSION
AC126962.5 GI:30521159
KEYWORDS
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)

```

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 222692)

Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alabrooks,S., Amin,A., Anguiano,D.,

Ayalabechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhaq,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Chen,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,

Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,

Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,

Geбреgeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gumaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Khan,Z., Kelly,S., Khan,Z., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louisaged,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nair,L.,
Nwakoelameh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,W., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanik,K.,
Valas,R., Vera,V., Villanueva,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,P., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE

JOURNAL

AUTHORS

REFERENCE

2 (bases 1 to 222692)

Unpublished

Direct Submission

Submitted (11-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 222692)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:23664494.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GQJ

Center clone name: CH230-132K18

----- Summary Statistics

Assembly program: Atlas 3.0

Consensus quality: 218983 bases at least Q40

Consensus quality: 220425 bases at least Q30

Consensus quality: 221525 bases at least Q20
 Estimated insert size: 229249; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 2 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

* 1 220835: contig of 220835 bp in length
 * 220836 220935: gap of unknown length
 * 220936 222692: contig of 1757 bp in length.

FEATURES

source

1..222692
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-132K18"
 1..1783
 /note="wgs contig"
 1834..3259
 /note="wgs contig"
 complement(218613..219445)
 /note="clone boundary
 clone end:Sp6
 site:EcORI
 end_sequence:BH321266"
 220836..220935
 /estimated_length=unknown

ORIGIN

Alignment Scores:
 Pred. No.: 1.07e+04 Length: 222692
 Score: 39.00 Matches: 8
 Percent Similarity: 88.9% Conservative: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 88.6% Indels: 0
 DB: 14 Gaps: 0

US-10-774-176-5 (1-9) x AC126962 (1-222692)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 |||||

Db 22931 TTCTCAGGAAACACGACGCTGTC 22957
 |||||

RESULT 30
 AC127173

LOCUS AC127173 228883 bp DNA linear ROD 30-JUL-2003
 DEFINITION Mus musculus strain C57BL/6J clone rp23-39k4 map 10, complete
 sequence.

ACCESSION AC127173

VERSION AC127173.17 GI:33342320

KEYWORDS HTG

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 228883)

Do.T. and Roe.B.A.

Mus musculus Chromosome 10 BAC Clone rp23-39k4

Unpublished

2 (bases 1 to 228883)

Do.T. and Roe.B.A.

Direct Submission

Submitted (14-JUL-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE

3 (bases 1 to 228883)

Do.T. and Roe.B.A.

AUTHORS

TITLE

JOURNAL

Submitted (24-APR-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

4 (bases 1 to 228883)

Do.T. and Roe.B.A.

AUTHORS

TITLE

JOURNAL

Submitted (30-JUL-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

COMMENT

On Jul 30, 2003 this sequence version replaced gi:30089776.

CENTER

INFORMATION

CENTER

CODE

UOKNOR

CENTER

ADDRESS

LOCATION

QUALIFIERS

SOURCE

1..228883

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/map="10"

/clone="rp23-39k4"

/note="This is one of two clones in the same well from
 rp23-39k4"

ORIGIN

ALIGNMENT

SCORES

PRED. NO.:

1.1e+04

LENGTH:

228883

SCORE:

39.00

MATCHES:

8

PERCENT SIMILARITY:

88.9%

CONSERVATIVE:

0

BEST LOCAL SIMILARITY:

88.9%

MISMATCHES:

1

QUERY MATCH:

88.6%

INDELS:

0

GAPS:

0

DB:

9

US-10-774-176-5 (1-9) x AC127173 (1-228883)

Qy

1

PheLeuThrGlyAsnGlnLeuAlaVal 9

|||||

Db

194288

TTCTCAGGAAATCACTGAAGTG 194314

RESULT 31

AC159326

LOCUS

AC159326

DEFINITION

Mus musculus 10 BAC RP23-45H6 (Roswell Park Cancer Institute
 (C57BL/6J Female) Mouse BAC Library) complete sequence.

ACCESSION

AC159326

VERSION

AC159326.9 GI:66792960

KEYWORDS

HTG

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 230689)

Muzny, D., Adams, C., Agbai II, O., Allen, C., Alsbrooks, S., Archer, P.,
 Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
 Berducci, A., Biswal, K., Blyth, P., Bonham, H., Buha, C., Burch, P.,
 Cadore, J., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
 Chacko, J., Chahrouh, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
 Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
 Curry, S., Dai, W., Davila, M.L., Davis, C., Davy-Carroll, L., De
 Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
 Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
 Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
 Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
 Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
 Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,
 Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeberlen, K.,

Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,
Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B.,
Howell, L.T., Huiyk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,
Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalafus, K., Kelly, S.,
Keys, T., Khan, Z., King, L., Kovar, C., Kowis, C., Lara, F.,
Leal, S., Lee, K., Lee, S., LeGall, P.I., Lemon, S., Lewis, L., Li, B.,
Li, Y., Li, Z., Linnell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D.,
London, P., Lopez, J., Lorensuikawa, L., Lozano, R., Luk, T., Madu, R.,
Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, E.,
McLellan, H., McPherson, J., Mercadado, C., Metzker, M.,
Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munidasa, M.,
Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Eastaugh, E.,
Nott, A., Nwakoelam, O., Obregon, M., Ochi-Okorie, C., Odeh, E.,
Okwuonu, G., Okwuonu, K., Parker, D., Pasternak, S., Patel, B.,
Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T.,
Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J.,
Rabata, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S.,
Rives, C., Rodriguez, F., Rojas, A., Ruiz, S.J., Sana, M., Sanders, W.,
Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,
Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R.,
Svatek, A., Taylor, E., Taylor, T., Thomas, N., Thorn, R., Thornton, R.,
Trejos, Z., Usmani, K., Vargo, C., Verduzco, D., Villasana, D., Virk, D.,
Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,
Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R.,
Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,
Zhang, Z., Zhou, J., Weinstock, G. and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (01-APR-2005) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 230689)
Worley, K.C.
Direct Submission
Submitted (19-MAY-2005) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 230689)
Worley, K.C.
Direct Submission
Submitted (28-MAY-2005) Human Genome Sequencing Center, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 28, 2005 this sequence version replaced gi:66275319.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by
Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui
Zhang.

Exon/intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Drafting Center Code: WIER

Contact: hgsc-help@bcm.tmc.edu.

Location/Qualifiers

1. .230689

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="10"

/clone="RP23-451H6"

FEATURES

source

misc_feature
complement(1..47110)
/note="overlaps bases 1..47110 of clone AC153548"
/function="clone overlap"
335..397
/rpt_family="TA)n"
repeat_region
763..800
/rpt_family="TGAA)n"
complement(1909..2073)
repeat_region
/rpt_family="MTE"
2637..2725
/rpt_family="B4A"
repeat_region
2650..2801
/rpt_family="B1F1"
repeat_region
4048..4190
/rpt_family="B4A"
repeat_region
4204..4243
/rpt_family="CATATA)n"
complement(4368..4405)
repeat_region
/rpt_family="RMER17C"
4406..4578
/rpt_family="GA)n"
repeat_region
complement(4579..4750)
repeat_region
/rpt_family="RMER17C"
4751..4841
/rpt_family="L1Md_F3"
complement(5306..5636)
repeat_region
/rpt_family="L1_Mm"
5635..6661
/rpt_family="L1Md_F2"
repeat_region
6662..6757
/rpt_family="GAAA)n"
complement(6764..6862)
repeat_region
/rpt_family="MYSERVIC_1"
complement(6889..7517)
repeat_region
/rpt_family="L1_Mm"
7518..8950
/rpt_family="Lx"
repeat_region
9097..9178
/rpt_family="B2_Mm2"
repeat_region
9557..9773
/rpt_family="Lx8"
repeat_region
11284..11359
/rpt_family="RLTR16"
11460..11574
/rpt_family="RLTR17"
11658..11714
/rpt_family="BGLII_B"
11848..12038
/rpt_family="GA-rich"
repeat_region
13371..13426
/rpt_family="AT-rich"
complement(13498..13768)
repeat_region
/rpt_family="MLT1B"
13957..13984
/rpt_family="AT-rich"
complement(15671..15759)
repeat_region
/rpt_family="MTE"
complement(15762..16416)
repeat_region
16417..16459
/rpt_family="MTE-int"
complement(16460..16837)
repeat_region
/rpt_family="MTE-int"
17097..17349
/rpt_family="Lx5"
repeat_region
17382..17959
/rpt_family="L1_Mus4"
complement(17960..18386)
repeat_region
/rpt_family="BGLII"
18387..20193
/rpt_family="L1_Mus4"
20213..20323
/rpt_family="Lx5"

```

repeat_region 20338..22096
/rpt family="Lx5"
repeat_region 22125..22186
/rpt family="Lx5"
repeat_region complement(22319..22524)
/rpt family="WTE-int"
repeat_region 22875..22955
/rpt family="RMER48"
repeat_region 23156..23455
/rpt family="ORRIC2"
repeat_region 23813..23958
/rpt family="B1_Mus2"
repeat_region 23959..23979
/rpt family="(A)n"
repeat_region 24115..24468
/rpt family="WTE"
repeat_region 24840..24880
/rpt family="(TG)n"
repeat_region complement(24906..24966)
/rpt family="TRNA-Ala-GCA"
repeat_region complement(25702..25779)
/rpt family="MIR"
repeat_region complement(26349..26494)
/rpt family="MLT1C"
repeat_region complement(27422..27562)
/rpt family="RSINE1"
repeat_region 27563..27602
/rpt family="AT-rich"

```

Alignment Scores:

```

Pred. No.: 1.11e+04 Length: 230689
Score: 39.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 88.6% Indels: 0
DB: 9 Gaps: 0

```

US-10-774-176-5 (1-9) x AC159326 (1-230689)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9

```

|||||
Db 98809 TTCTCTACTGGAATCACTGAAGTG 98835

```

RESULT 32

AC106580/c

LOCUS AC106580 234843 bp DNA linear HTG 13-MAY-2003

DEFINITION Rattus norvegicus clone CH230-20206, *** SEQUENCING IN PROGRESS

*** 6 unordered pieces.

ACCESSION AC106580

VERSION AC106580.4 GI:30579811

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus

1 (bases 1 to 234843)

Muzny, D., Marle, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalbechini, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Dexamco, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gabregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Loulseged, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smajs, D., Sneed, A., Sodergren, E., Song, X., Sorelle, R., Soosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D., Holt, R., Smith, H., O., Weinstock, G., and Gibbs, R. A.

TITLE Direct Submission

REFERENCE 2 (bases 1 to 234843)

Worley, K. C.

Direct Submission

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 234843)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23664780.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLKG

Center clone name: CH230-20206

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 215513 bases at least Q40

Consensus quality: 218526 bases at least Q30

Consensus quality: 220428 bases at least Q20

Estimated insert size: 232032; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 227313: contig of 227313 bp in length
* 227314 227413: gap of unknown length
* 227414 228546: contig of 1133 bp in length
* 228547 228646: gap of unknown length
* 228647 230164: contig of 1518 bp in length
* 230165 230264: gap of unknown length
* 230265 231593: contig of 1329 bp in length
* 231594 231693: gap of unknown length
* 231694 232948: contig of 1255 bp in length
* 232949 233048: gap of unknown length
* 233049 234843: contig of 1795 bp in length.

FEATURES
source
1..234843
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clones="CH230-20206"
66562..66637
/note="clone boundary
clone_end:sp6
site:ECORI
end sequence:BH335327"
227314..227413
/estimated_length=unknown
228547..228646
/estimated_length=unknown
230165..230264
/estimated_length=unknown
231594..231693
/estimated_length=unknown
232949..233048
/estimated_length=unknown

ORIGIN

Alignment Scores:
Pred. No.: 1.13e+04 Length: 234843
Score: 39.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 88.6% Indels: 0
DB: 14 Gaps: 0

US-10-774-176-5 (1-9) x AC106580 (1-234843)

QY 1 PheLeuThrGlyAenGlnLeuAlaVal 9
|||||

Db 62928 TTCTCAGCAACACGACAGCTGTC 62902
|||||

RESULT 33
AC128267
LOCUS
DEFINITION Rattus norvegicus clone CH230-46018, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
AC128267
VERSION AC128267.2 GI:23196313
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

REFERENCES
AUTHORS

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

1 (bases 1 to 259795)

Muzny D, Maric, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalbechchi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biawalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, B., Louisleged, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapus, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankech, C., Plopper, F., Poindester, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajz, D., Sneed, A., Sodergren, E., Song, X. Z., Sorell, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanai, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 259795)

Worley, K. C.

Direct Submission

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 259795)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Sep 19, 2002 this sequence version replaced gi:21908887. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYCL
Center clone name: CH230-46018
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 234408 bases at least Q40
Consensus quality: 237198 bases at least Q30
Consensus quality: 238865 bases at least Q20
Estimated insert size: 256961; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 255963: contig of 255963 bp in length
* 255964 256063: gap of unknown length
* 256064 259795: contig of 3732 bp in length.
FEATURES             Location/Qualifiers
     source            1..259795
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-46018"
     misc_feature      1..1625
                        /note="wgs_end_extension
                        clone_end:T7"
                        7261..8603
     misc_feature      /note="wgs_end_extension
                        clone_end:T7"
     misc_feature      complement(9310..9645)
                        /note="clone_boundary
                        clone_end:T7
                        site:EcoRI
     misc_feature      end_sequence:BH331998"
                        complement(246071..246516)
                        /note="clone_boundary
                        clone_end:Sp6
                        site:EcoRI
     misc_feature      end_sequence:BH331999"
                        250515..253075
                        /note="wgs_end_extension
                        clone_end:Sp6"
     misc_feature      254334..255963
                        /note="wgs_end_extension
                        clone_end:Sp6"
     gap              255964..256063
                        /estimated_length=unknown
     misc_feature      256064..257208
                        /note="wgs_end_extension
                        clone_end:Sp6"
     misc_feature      257209..259795
                        /note="wgs_end_extension
                        clone_end:Sp6"
ORIGIN
Alignment Scores:
Pred. No.: 1.24e+04 Length: 259795
Score: 39.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1

```

```

Query Match: 88.6% Indels: 0
DB: 14 Gaps: 0
US-10-774-176-5 (1-9) x AC128267 (1-259795)
QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
   |||||
Db 227753 TTTCTCACAGGAACCCAGACAGCTGTC 227779
   |||||

RESULT 34
AL807132 365 bp DNA linear STS 11-JUN-2003
LOCUS Arabidopsis thaliana transposon insertion STS SM_3.19725, sequence
DEFINITION tagged site.
ACCESSION AL807132
VERSION AL807132.1 GI:21622180
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B.,
          Murphy,G., Langham,S., LeGrys,C., Jones,J.D.G. and Bevan,M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 365)
AUTHORS Clarke,J.H.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2002) Clarke J.H., John Innes Centre, Colney
          Lane, Norwich, NR4 7UJ, UK
COMMENT At denotes an activation tag dissociation transposon within a
          single line, Et an enhancer trap dissociation transposon, Gt a gene
          trap dissociation transposon, Mt a mis-expression enhancer trap
          dissociation transposon, SM a defective suppressor mutator
          transposon. _3 denotes a sequence derived from the 3'end of the
          transposon, _5 denotes a sequence derived from the 5'end of the
          transposon BBSRC GARNET, ATIS project
          On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock
          code: N108624.
FEATURES             Location/Qualifiers
     source            1..365
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /variety="Columbia-0 NASC stock code N1092"
                        /db_xref="taxon:3702"
                        /clone="AC011809"
                        /note="Derived from superpool 18.11 NASC code N40810"
     STS               1..365
                        /standard_name="SM_3.19725"
ORIGIN
Alignment Scores:
Pred. No.: 38.4 Length: 365
Score: 38.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 86.4% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-5 (1-9) x AL807132 (1-365)
QY 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
   |||||
Db 86 TTTTACAGGCAATCATATTAGCCGTC 112
   |||||

RESULT 35
AR347195
LOCUS Arabidopsis thaliana transposon insertion STS PAT 17-AUG-2003
DEFINITION Sequence 1806 from patent US 6583275.
ACCESSION AR347195
VERSION AR347195.1 GI:33744240
KEYWORDS

```

```

SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 972)
AUTHORS     Doucette-Stamm, L.A. and Bush, D.
TITLE       Nucleic acid sequences and expression system relating to
            Enterococcus faecium for diagnostics and therapeutics
JOURNAL     Patent: US 658375-A 1806 24-JUN-2003;
            Genome Therapeutics Corporation; Waltham, MA
FEATURES    Location/Qualifiers
            source
              1..972
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      98.3      Length:      972
Score:          38.00     Matches:      8
Percent Similarity: 88.9%   Conservative: 0
Best Local Similarity: 88.9% Mismatches:  1
Query Match:    86.4%     Indels:      0
DB:             6        Gaps:          0

US-10-774-176-5 (1-9) x AR347195 (1-972)

Qy      1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db      841 TTTCTGACAGGAGCAGCTCGCAGTA 867

RESULT 36
BD249732
LOCUS    BD249732                      1281 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249732
VERSION    BD249732.1 GI:33059502
KEYWORDS   JP 2002530060-A/2.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1281)
AUTHORS     Carroll, M.W. and Myers, K.A.
TITLE       Polypeptide
JOURNAL     Patent: JP 2002530060-A 2 17-SEP-2002;
            OXFORD BIOMEDICA LTD
COMMENT     OS Mus musculus (mouse)
            PN JP 2002530060-A/2
            PD 17-SEP-2002
            PF 18-NOV-1999 JP 2000582415
            PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
            30-JUL-1999 GB 9917995.4
            PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
            PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K14/065,
            PC C07K19/00,
            PC C12N15/00
            CC Polypeptide
            FH Key
            FT source
              1..1281
                Location/Qualifiers
                /organism="Mus musculus (mouse)"

FEATURES    source
            1..1281
              Location/Qualifiers
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      128      Length:      1281
Score:          38.00     Matches:      7
Percent Similarity: 88.9%   Conservative:  1
Best Local Similarity: 77.8% Mismatches:  1
Query Match:    86.4%     Indels:      0

US-10-774-176-5 (1-9) x AX025012 (1-1281)

Qy      1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db      289 TTTCTTACCGCAACACGATGACCGTG 315

RESULT 37
AX025012
LOCUS    AX025012                      1281 bp      DNA      linear      PAT 15-SEP-2000
DEFINITION Sequence 2 from Patent WO0029428.
ACCESSION AX025012
VERSION    AX025012.1 GI:10184933
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Carroll, M.W. and Myers, K.A.
TITLE       Polypeptide
JOURNAL     Patent: WO 0029428-A 2 25-MAY-2000;
            CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
            BIOMEDICA LTD (GB)
FEATURES    Location/Qualifiers
            source
              1..1281
                /organism="Mus musculus"
                /mol_type="unassigned DNA"
                /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      128      Length:      1281
Score:          38.00     Matches:      7
Percent Similarity: 88.9%   Conservative:  1
Best Local Similarity: 77.8% Mismatches:  1
Query Match:    86.4%     Indels:      0
DB:             6        Gaps:          0

US-10-774-176-5 (1-9) x AX025012 (1-1281)

Qy      1 PheLeuThrGlyAsnGlnLeuAlaVal 9
Db      289 TTTCTTACCGCAACACGATGACCGTG 315

RESULT 38
AX316087
LOCUS    AX316087                      1281 bp      DNA      linear      PAT 14-DEC-2001
DEFINITION Sequence 2 from Patent EP1160323.
ACCESSION AX316087
VERSION    AX316087.1 GI:17899279
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1
AUTHORS     Carroll, M.W. and Myers, K.A.
TITLE       St4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL     Patent: EP 1160323-A 2 05-DEC-2001;
            Oxford Biomedica (UK) Limited (GB)
FEATURES    Location/Qualifiers
            source
              1..1281
                /organism="Mus musculus"
                /mol_type="unassigned DNA"
                /db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.:      128      Length:      1281
Score:          38.00     Matches:      7
Percent Similarity: 88.9%   Conservative:  1
Best Local Similarity: 77.8% Mismatches:  1
Query Match:    86.4%     Indels:      0

```

Score: 38.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 86.4% Indels: 0
 Db: 6 Gaps: 0

US-10-774-176-5 (1-9) x AX316087 (1-1281)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 |||||
 Db 289 TTCCTTACCGCAACCATGACCGTG 315

RESULT 39
 AF063939
 LOCUS AF063939
 DEFINITION Rattus norvegicus 574 oncofetal antigen homolog (574) mRNA, ROD 01-JAN-2000 complete cds.

ACCESSION AF063939

VERSION AF063939.1 GI:6650211

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 2333)

AUTHORS Ninkina, N. and Buchman, V. I.

TITLE Structure and expression of the rat 574 gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2333)

AUTHORS Buchman, V. I.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-1998) School of Biomedical Sciences, University of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS, UK

FEATURES

source

Location/Qualifiers

1..2333

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/tissue_type="cerebellum"

/dev_stage="newborn"

1..2333

/gene="574"

1..363

/gene="574"

364..1644

/gene="574"

/codon_start=1

/product="574 oncofetal antigen homolog"

/protein_id="AAF21770.1"

/db_xref="GI:6650212"

/translation="WPGAGSRGPSAGDGRRLRLALVLGWSASAPSSLPSSSTS

PAAFIASSAQPPPAERCPAACESEARTVKVNRNLLLEVPADLPVPVRLNLFILGNQ

MTVLPAGAPQPPPLADLAVNLGNLKEVGAPFHLPLGLRLDLGHPNPLNLSAF

TPAGNSVFTPLLEILNIHVPPQKQNGFEGMVAFEGMAALRSLALRGL

HLLELASHFLYLPRLDLQPLSLKHLDRNNSLVLTVAFPNLTLESLHLEDONL

KVLNRSLTAEQGLAHVRVFLDNNPVCDCYMDVMVSLKETEVVPDKARLTCAPEK

MNRGLLDLTSDLDLDCATLPQSLQTSVFVIGIVIALGAIFLLVLYLNKGIKKWMH

NIRDAQRDHMGHYRIYNADPSLTNLSNSDV"

1645..2333

/gene="574"

2315..2320

/gene="574"

polyA_signal

ORIGIN

Alignment Scores:

Pred. No.: 228 Length: 2333

Score: 38.00 Matches: 7

Percent Similarity: 88.9% Conservative: 1

Best Local Similarity: 77.8% Mismatches: 1

Query Match: 86.4% Indels: 0

Db: 9 Gaps: 0

US-10-774-176-5 (1-9) x AF063939 (1-2333)

Qy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
 |||||
 Db 652 TTCCTCACTGGCAACCATGACCGTG 678

RESULT 40

BC087011

LOCUS BC087011

DEFINITION Rattus norvegicus trophoblast glycoprotein, mRNA (cDNA clone

MGC:93332 IMAGE:7193411), complete cds.

ACCESSION BC087011

VERSION BC087011.1 GI:56268819

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 2361)

AUTHORS Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,

Klausner, R. D., Collins, P. S., Wagner, K. H., Shenmen, C. M., Schuler, G. D.,

Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K.,

Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L.,

Stapleton, M., Soares, M. B., Bonaldo, M. F., Cabavant, T. L.,

Scheetz, T. B., Brownstein, M. J., Usdin, T. B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J.,

Abramson, R. D., Mullaby, S. J., Bosak, S. A., McGowan, P. J.,

McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S.,

Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W.,

Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A.,

Fahay, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y.,

Bouffard, G. G., Blakeley, R. W., Touchman, J. W., Green, E. D.,

Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M.,

Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smalish, D. E.,

Schnerch, A., Schein, J. E., Jones, S. J., and Marra, M. A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2361)

REFERENCE Direct MGC Project.

AUTHORS Direct Submission

TITLE Direct Submission

JOURNAL Direct Submission

Submitted (02-DEC-2004) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Howard Jacobs

cDNA Library Preparation: Express Genomics

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 186 Row: 0 Column: 24

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

source

1..2361

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

gene	/clone="MGC:93332 IMAGE:7193411"	
	/tissue_type="Heart, rat (Brown Norway)"	
	/clone_lib="NIH_MGC_234"	
	/lab_host="DH10B"	
	/note="Vector: pExpress1"	
	1..2361	
	/gene="Tpbg"	
	/note="synonym: 5T4"	
	/db_xref="GeneID:83684"	
	/db_xref="RGD:621453"	
CDS	364..1644	
	/gene="Tpbg"	
	/codon_start=1	
	/product="Tpbg protein"	
	/protein_id="AAH87011.1"	
	/db_xref="GI:56268820"	
	/db_xref="GeneID:83684"	
	/db_xref="RGD:621453"	
	/translation="MPGAGSGPSAGDGRRLRLARLALVLLGWVSASAPSSLPSSSTS	
	PAAFLASGSAQPPPAECRPAACESEARTVKCVNRNLLLEVPADLPVVRNLFITGNO	
ORIGIN	MTVLPAQAPARQPLADLAVNLNLSGNHLKEVGAGAFELHGLRLDLSHNPILNLSAF	
	TFAGSNVSTPSPLEILNHIVPDQRNGSFEGVAFEGVAAALRSGLALRGL	
	HLELASNHFLYPRDLDDQLPSLKHLDLRNLSVLTYSFASFNLTLESJLSDNAL	
	KVLHNSLAQWQGLAHVRFLDNNPWCDCYMDMVSFKETEVVPDKARLTCAPEK	
	MNRNGLDLSDDLDCDATALPQSLQTSVFLGIVLALIGALFLVLYLNRKGIKKMMH	
	NIRDACRDHMEGYHYRINADPRLTNLSSNDV"	
	642..1262	
	/gene="Tpbg"	
	/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)	
	misc_feature	
FEATURES	source	
	1..2423	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="FVB/N"	
	/db_xref="taxon:10090"	
	/clone="MGC:68145 IMAGE:5353871"	
	/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating	
	ductal Carcinoma. 5 month old virgin mouse."	
	/clone_lib="NCI CGAP_Mam6"	
gene	/lab_host="DH10B"	
	/note="Vector: pCMV-SPORT6"	
	1..2423	
	/gene="Tpbg"	
	/note="synonym: 5T4"	
	/db_xref="GeneID:21983"	
	/db_xref="MGI:1341264"	
	402..1682	
	/gene="Tpbg"	
	/codon_start=1	
CDS	/product="Tpbg protein"	
	/protein_id="AAH58198.1"	
	/db_xref="GI:34849574"	
	/db_xref="GeneID:21983"	
	/db_xref="MGI:1341264"	
	/translation="MPGAGSGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSSTS	
	PAAFLASGSAQPPPAECRPAACESEARTVKCVNRNLLLEVPADLPVVRNLFITGNO	
	MTVLPAQAPARQPLADLAVNLNLSGNHLKEVGAGAFELHGLRLDLSHNPILNLSAF	
	TFAGSNVSTPSPLEILNHIVPDQRNGSFEGVAFEGVAAALRSGLALRGL	
	HLELASNHFLYPRDLDDQLPSLKHLDLRNLSVLTYSFASFNLTLESJLSDNAL	
AUTHORS	KVLHNSLAQWQGLAHVRFLDNNPWCDCYMDMVSFKETEVVPDKARLTCAPEK	
	MNRNGLDLSDDLDCDATALPQSLQTSVFLGIVLALIGALFLVLYLNRKGIKKMMH	
	NIRDACRDHMEGYHYRINADPRLTNLSSNDV"	
	642..1262	
	/gene="Tpbg"	
	/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)	
	misc_feature	
	10-774-176-5 (1-9) x BC087011 (1-2361)	
	1 PheLeuThrGlyAanGlnLeuAlaVal 9	
652 TTCTCCTGCGCAACGATGACCGTG 678		
BUILT 41	BC058198	
	2423 bp mRNA linear ROD 21-OCT-2003	
	Mus musculus trophoblast glycoprotein, mRNA (CDNA clone MGC:68145	
	IMAGE:5353871), complete cds.	
	BC058198	
	GI:34849573	
	MGC.	
	Mus musculus (house mouse)	
	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
ORIGIN	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;	
	Sciurognathi; Muridea; Muridae; Murinae; Mus.	
	1 (bases 1 to 2423)	
	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,	
	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,	
	Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,	
	Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,	
	Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,	
	Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,	
	Sheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,	
REFERENCE	Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,	
	Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,	
	McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,	
	Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,	
	Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,	
	Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,	
	Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,	
	Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,	
	Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,	
	Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smailus, D.E.,	

```

RESULT 43
AX961914
LOCUS      AX961914
DEFINITION Mus musculus (house mouse)
ACCESSION GI:40881324
VERSION    1
KEYWORDS   Mus musculus
SOURCE     Mus musculus
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS    Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
TITLE       Stat6 activation gene
JOURNAL     Patent: WO 03104277-A 125 18-DEC-2003;
            Asahi Kasei Kabushiki Kaisha (JP)

FEATURES
source     1..2557
            /organism="Mus musculus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:10090"
            556_..1836
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAF06466.1"
            /db_xref="GI:40881325"
            /translation="MPGAGSRGPSAGDGRLELARLALVLLGWVSASAPSSVPSSSTSP
            PADPLASGAQPPAERCPAAECESARTKVCNRRNLLEVADLPYPYVRNFLGTNQSAF
            MTVPLPAGAFARPQLADLRLNALNSGNHLKEVCAGAEHUPGLRRLDLSHNPFLTNLSAF
            VFGANSVASPSPEELILNIHVPPEDQRONGSFEGWAFEGMVAAALRSGLRALRG
            TLELASNHFWLFPRDLLAQPLSLRYLDLRNNLSVLITVASFNLTLESLSHLIEDNAL
            KLHNSTLAHEWGCLAHVKVFLDNRPWCYMDMVAWLKETEVDPDKARLTCAFPBPK
            MNRGILLDNSDLDCDAVLPOSLOTSYVPLGLVLALIGAILFLVLYLNKKGIKKWMI
            NRDACRDMEGHYYEYNADPLNLSSNDV"

ORIGIN
Alignment Scores:
Pred. No.:      249          Length:      2557
Score:           38.00        Matches:      7
Percent Similarity: 88.9%    Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match:      86.4%     Indels:      0
DB:               6         Gaps:        0

US-10-774-176-5 (1-9) x AX961914 (1-2557)

Qy      1 PheLeuThrGlyAsnGlnLeuAlaVal 9
          |||||
Db      844 TTCCTTACC CGCAAC CAGATGCCGTG 870

RESULT 44
MMU012160
LOCUS      MMU012160
DEFINITION Mus musculus 574 oncofetal trophoblast glycoprotein gene.
ACCESSION J012160
VERSION    1
KEYWORDS   Mus musculus
SOURCE     Mus musculus
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS    King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE       Organisation of the mouse and human 574 oncofoetal leucine-rich
            glycoprotein genes and expression in foetal and adult murine
            tissues
JOURNAL     Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED     10366710
REFERENCE  2 (bases 1 to 7942)
AUTHORS    Myers, K.A.

```


TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
 Institute for Cancer Research, Christie Hospital, Wilmslow Road,
 Manchester, M20 9BX, UK

FEATURES
 source

1. 7942
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129/Sv"
 /db_xref="taxon:10090"
 /clone_lib="Lambda Dash"
 3108. 3113
 /bound_moiety="Spl"
 3114. 3119
 /bound_moiety="Spl"
 3124. 5779
 /gene="574"
 3124. 3151
 /gene="574"
 3152. 3450
 /gene="574"
 3451. 5779
 /gene="574"
 3779. 5059
 /gene="574"
 /codon_start=1
 /product="574 oncofetal trophoblast glycoprotein"
 /protein_id="CAA09931.1"
 /db_xref="GI:3805949"
 /db_xref="GOA:Q920L0"
 /db_xref="InterPro:IPR000372"
 /db_xref="InterPro:IPR000483"
 /db_xref="InterPro:IPR001611"
 /db_xref="InterPro:IPR003591"
 /db_xref="MGI:1341264"
 /db_xref="UniProt/TREMBL:Q920L0"
 /translation="MPGSGSGPSAGDGLRLARLALVLLGMVSASAPSSVPSSSTS
 PADPLASGAPPPAPKCPACRCESEARTVKCVNRLLEVPADLPFYVRLFTGNQ
 MVLPAFAFAPQPLADLELNLSGNHLKEVCAGAFELHPLGLRLDLNHPNLTLSAF
 VFAGSNASVAPSLLEILNHIYPEDQRQNGFEGVAFEGVAAALRSLGLRGL
 TRLEASNLHPLFLPRDLIAQLPSRLYLDLRNLSLVLTYSFRNLTHLESRLHLDNAL
 KVLHNSLAWQGLARVKVFLDNNPWVCDVMADVAWLKETEVPVDPKARUTCAPEK
 MNRGLDLNSSLDDCDVLPQSLQTSVFLGIVLALGILFLVLVLYNRGKIKWHE
 NIRDACRDHSGHYHYRINADPRLTNSSDV"
 3779. 3865
 /gene="574"
 3866. 5056
 /gene="574"
 /product="574 oncofetal trophoblast glycoprotein"
 5713. 5718
 /gene="574"
 5759. 5764
 /gene="574"

ORIGIN

Alignment Scores:
 Pred. No.: 740 Length: 7942
 Score: 38.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 86.4% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-5 (1-9) x MMU012160 (1-7942)

Qy 1 PheLeuThrGlyAenGlnLeuAlaVal 9

Db 4067 TTCCTTCCGCAACACGATGACCGTG 4093

RESULT 45

AB001073

LOCUS

DEFINITION

Archaeoglobus fulgidus DSM 4304 section 34 of 172 of the complete

linear BCT 17-MAR-2003

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

genome.
 AE001073 AE000782
 AE001073.1 GI:2689396
 Archaeoglobus fulgidus DSM 4304
 Archaeoglobus fulgidus DSM 4304
 Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 Archaeoglobaceae; Archaeoglobus.

REFERENCE
AUTHORS

1 (bases 1 to 15951)
 Klenk,H.P., Clayton,R.A., Tomb,J., White,O., Nelson,K.E.,
 Ketchum,K.A., Dodson,R.J., Gwinn,M., Hickey,E.K., Peterson,J.D.,
 Richardson,D.L., Kerlavage,A.R., Graham,D.E., Kyrpides,N.C.,
 Fleischmann,R.D., Quackenbush,J., Lee,N.H., Sutton,G.G., Gill,S.,
 Kirkness,E.F., Dougherty,B.A., McKenney,K., Adams,M.D., Loftus,B.,
 Peterson,S., Reich,C.I., McNeil,L.K., Badger,J.H., Glodek,A.,
 Zhou,L., Overbeek,R., Gocayne,J.D., Weidman,J.F., McDonald,L.,
 Utterback,T., Cotton,M.D., Spriggs,T., Artlich,P., Kaine,B.P.,
 Sykes,S.M., Sadow,P.W., D'Andrea,K.P., Bowman,C., Fujii,C.,
 Garland,S.A., Mason,T.M., Olsen,G.J., Fraser,C.M., Smith,H.O.,
 Woese,C.R. and Venter,J.C.

TITLE
JOURNAL
PUBMED
REFERENCES
AUTHORS

The complete genome sequence of the hyperthermophilic,
 sulphate-reducing archaeon Archaeoglobus fulgidus
 Nature 390 (6658), 364-370 (1997)
 9389475
 2 (bases 1 to 15951)
 Klenk,H.P., Clayton,R.A., Tomb,J.-F., White,O., Nelson,K.E.,
 Ketchum,K.A., Dodson,R.J., Gwinn,M., Hickey,E.K., Peterson,J.D.,
 Richardson,D.L., Kerlavage,A.R., Graham,D.E., Kyrpides,N.C.,
 Fleischmann,R.D., Quackenbush,J., Lee,N.H., Sutton,G.G., Gill,S.,
 Kirkness,E.F., Dougherty,B.A., McKenney,K., Adams,M.D., Loftus,B.,
 Peterson,S., Reich,C.I., McNeil,L.K., Badger,J.H., Glodek,A.,
 Zhou,L., Overbeek,R., Gocayne,J.D., Weidman,J.F., McDonald,L.,
 Utterback,T., Cotton,M.D., Spriggs,T., Artlich,P., Kaine,B.P.,
 Sykes,S.M., Sadow,P.W., D'Andrea,K.P., Bowman,C., Fujii,C.,
 Garland,S.A., Mason,T.M., Olsen,G.J., Fraser,C.M., Smith,H.O.,
 Woese,C.R. and Venter,J.C.

TITLE
JOURNAL

Submitted (15-DEC-1997) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA

REMARK

In order to show the genes in ascending order on the genome, the
 origin of this version has been moved by TIGR to position 2093570
 of the original version and the opposite strand is shown from the
 original version.

On Dec 16, 1997 this sequence version replaced gi:2650166.

COMMENT
FEATURES
 source

1. 15951
 /organism="Archaeoglobus fulgidus DSM 4304"
 /mol_type="genomic DNA"
 /db_xref="taxon:224325"
 81. 2396
 /gene="AF0448"
 81. 2396
 /gene="AF0448"
 /note="similar to PID:1001216 PID:1001267 percent
 identity: 26.09; identified by sequence similarity;
 putative"
 /codon_start=1
 /transl_table=11
 /product="signal-transducing histidine kinase, putative"
 /protein_id="AAB90784.1"
 /db_xref="GI:2650176"

gene

CDS

/translation="MDSSEIKIWEWFKEVSKFSPBLRVVVVKERIKGYIALLSGE
 LRKMLVYIDDELIERIIRSGKPDVVRVRLMELFANKLSEYEFEPKVAKNFAFPFVAL
 YAVAVGAEKDKTKVPAASGNDYVAVIKVTWNSDSETPGPGSSVRLKRVVRE
 TDADESKPFEELKSGFSRVAALPLHVENVIGSLAIYAPDPAPFEEELLEEL
 AELSALASLREKKEMLKELYLVVNTGTALLVIENGKIVPANKEVEKLSYGR
 BELGKPTILLPEEDREKILRIHKLRLADPEAPRQYEIKTVDFGRIYRIITATV
 IPGTKRVVVSMLDITEAKMERKLSEERYRAVFPENSLGIVTASVDTIDCNAA
 LRMFMRREDVIGKTKBLGPDAAKLPIKKEFYRGLSEEPFLEVINIGGKIL
 KVPFVLLKKGVPVFLNIIEITETKMAERKLRLDALESQILRLVDLGIIRGNLRD
 MLLSVKVRKIGCDVTAIEVFGENGFPVSDPEIGRSVLEKIDAEKEVVNLEKDD
 LNEVELLKGGPARVYISPLIARDEKGVIFAAKSGIEGKIDFLFKTAGQLATAIH


```

this BAC (gb|AC011809), 24 bp before the predicted start
codon."
/codon_start=1
/product="Calcium-dependent protein kinase 1"
/protein_id="AAF27092.1"
/db_xref="GI:6730697"
/translation="MGNCACVCPDSKESKPSKPKPNRDRKLNPFAGDFTSRPAPI
RVLDKVPMSQTOISDKYILRGELRGEGITVLTCTDRETHEALACKSISKRLRTA
VDIEDVREVAIMSTLPEHNVNKLKASYEDNENVHLMELCEGELPDRIVARGHYT
ERAASAVARTIAEVVMCHSNGVHRDLKPENFLPANKKSPKLAIDFGLSVFPKPG
DKFTIIVGSPYTMALPEVKRDYDGVDMVMSAGVIYIILCGVFPFMASTEGGVALAIL
IVRDLKQPSMMNRFPKKVLRIABHLSIQBEVITKMFSLMDMDCKDKITYPELKG
LQKYSQLGEPIKMLMEVADVQNGFLDYGFVAIVIIHLOKIENDELFLKLAEMFEDK
DGSYIHELDELREALABELGDSVLSIDINREVDTKDGRINDEYFVTMMKAGTDMR
KASQYRSRFRKFSISINIMKDGSLHLHDHALTGQIVPV"
/complement(6629..9070)
/gene="F6A14.2"
/complement(join(6629..7529,7607..8160,8258..8475,
8980..9070))
/gene="F6A14.2"
/notice="Similar to peptide transport proteins"
/codon_start=1
/protein_id="AAF27093.1"
/db_xref="GI:6730698"
/translation="MEVEKTEKNIAEDDESILYRGWKVMPFIIGNETPEKLGIVGSS
SNLYITLVFNKMSITAAKVNIYGGTSNFGITVAAPLCSYGYRYKILGSPAMTACF
LGSVAMDLTAIVHLPQAQKISGVCNFSIQIMPLAGAMVLLVIGAGIRPCNL
PFGADQDPKTKGKRGIESPFNFFTPAQVSLTVVQSNVSWSLGAIAPAI
LMLGCIIFPAGSKLYVKVKGSPHISITRIVVIAIKRRLKPGPNELNYIASDF
KNSKLHTEOPFLDKSAIQODDKLNKSPDAMKLCMQOVEEVKCVIRLVPWL
SAALFYLAIOQYTTIFQSLQSDRRRQPGSQIPAGSYTVFLMLGNTIFIPIDRVL
VFILRYTGRDGIQOLQVAGAGFLCITSMVSAIVQYRKVKVLTITGLAPKRG
AISMGMWLTIPQVLMGIALDAGVGQMEFYKQFPENMFSPAGSLYCYGIGLASYL
STFLSVAHDTTFEGSGWLPEDLNKGRLEYFVFLVAGMTNLALAYFLNLSHWTRYK
DVVAKDKMDKTSAEFDKVS"
/complement(10694..14124)
/gene="F6A14.3"
/complement(join(10694..10741,10837..10929,11012..11109,
11244..11328,11532..11642,11809..11866,11962..12080,
12177..12245,12346..12420,12498..12644,12724..12840,
12920..13037,13131..13366,13960..14124))
/gene="F6A14.3"
/notice="Putative isochorismate synthase"
/codon_start=1
/protein_id="AAF27094.1"
/db_xref="GI:6730699"
/translation="MNGCEADHKAPLGTVTRTLSTVSPAAATERLITAVSDLKSQP
PPSSGIVRLQVPIEQKIGALDWLHAQNEILPRSPRSDSGRDLLQDPSSDNGSS
DHNPSVAGISAVFRDLDPFDDHDSIRPLSSKSLRAYGGLPDPFGKIAVE
WEHFGSPYFTVPQVDFGSSMLAATVAMNLSLWTLENAIEALQETMLQVSSVIM
RLRESLGVIVVSKNHVPSGAYTPAVNNALEIKDHSPLSKVVLARSRIITDID
PIAWLARQEGODAYQFCLOPGCAPIGNITPERLPHRKLHGVCSKLARVROGL
SKVEMLEBDLLTSPKDDLEFSIVRENREKLKICDRVVVKPHKSVALARVOHL
SOLQALQRDEDEFNLTALPTAVCCPVEARLLIKQIESPRGMVAGPIGFGG
GESFSGVIRSAIVKGLIYAGTGVISGNSPSSSEWNELELKISQFTKSLHESAL
QPIN"
/complement(18661..20376)
/gene="F6A14.5"
/complement(join(18661..19422,19510..19623,19837..20376))
/gene="F6A14.5"
/notice="Hypothetical protein"
/codon_start=1
/protein_id="AAF27095.1"
/db_xref="GI:6730700"
/translation="NDEAKENRKLKSLIKKDPDILQTOYNOLMAKNHPTKFSQ
KXHQDQGEDSDREKVNREKELVSLISGRRLNSENVPSSNKEKNKDYEEAGSDNYD
DNEKSIQGLSMGIEYKALSNPNEKLEIDHNOFTWLSISNNKILRSQNSFGKNDQ
DHEDEILPQNLVKTRVSNMDQWRYKQKAKGNPCPRAYRTRITAAFCPVKRG
VQSCEDMSILITVYEGTHNPLPMSATAMASATSAASMLLSGSSSSAAADLHLG
NRLSLGNNITPKPTKTHPLQSPSSGHPTVLDLTTSSSSQOPFLSMLNRFSPGNVS
RNSYPSNLTNFSNNTWNTLMWGGGNSDOYRAAYGNINTHQOSPYHKIIOITRTAGS
SDFPGRSSSHSQINLDHIGKNIISHQVPSLPASTIKAITTDPFSQSALATLSS

```

```

VERSION AC011809.2 GI:6579253
KEYWORDS HTG
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE
AUTHORS Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 108767)
Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.P.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.S.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Unpublished
2 (bases 1 to 108767)
Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.P.,
Altati, H., Nguyen, M., Lam, B., Southwick, A., Bei, Q., Buehler, E.,
Chin, C., Chlou, J., Choi, E., Dunn, P., Gonzalez, A., Howng, B., Kim, C.,
Koo, T., Lee, J.M., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S.,
Mukharasy, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P.,
Thavari, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J.,
Theologis, A. and Davis, R.W.
Direct Submission
Submitted (15-OCT-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
3 (bases 1 to 108767)
Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.P.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
Submitted (15-DEC-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
4 (bases 1 to 108767)
Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.P.,
Altati, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologis, A.
and Davis, R.W.
Direct Submission
Submitted (22-JAN-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
On Dec 15, 1999 this sequence version replaced gi:6041764.
e-mail for correspondence: arabsequence.stanford.edu
Genes with similarity to proteins in the databases are described
as 'putative', '-like' or 'similar to'. Genes that have EST
similarity but no significant protein similarity are described as
'unknown proteins'. Genes that are annotated based only on gene
prediction software are described as 'hypothetical proteins'. The
software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
http://genome.stanford.edu/~chris/GENSCANW.html), Fexa (V.Solovyev
& A.Salamov, Sanger Centre, http://genomic.sanger.ac.uk/), and
NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/NetPlantGene.html).
Location/Qualifiers
1. 108767
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F6A14"
1134..6402
/gene="F6A14.1"
join(4134..4785,4919..5062,5155..5307,5397..5512,
5621..5788,5981..6111,6229..6402)
/gene="F6A14.1"
/notice="Calcium-dependent protein kinase 1, gb|846283, the
cloned cDNA for this protein starts at nucleotide 4160 of

```

```

gene
  IMGDGDKIDHNVTREAREKSP"
  23743..24942
  /gene="F6A14.6"
CDS
  23743..24942
  /gene="F6A14.6"
  /note="Unknown protein; Location of ESTs TASG082,
  gb|Z18037 and 158G12T7, gb|AA720219"
  /codon_start=1
  /protein_id="AAF27096.1"
  /db_xref="GI:6730701"
  /translation="MTLEAEISANAVANQKVDQSQSDYPIPLSHDGIIFANLKPJENPN
  LGTLNPTISGMGIDSEVEIDLGKFSKLRKLKDYNGFVKMGKPLFKRIG
  EKVGISEAESVLDQAGSEIQILVEKNGFLMGDVGVLGKDCISLRLMELVELLI
  SNSLDHSYSVLNLEKORSDLLCVVKEASDLGATELLSILKFLCPKSEKALIT
  MAKVREESQAMLAIEKVNTELSKSKVAEASILLVAHDGFSLELCUHLIAS
  RNVDVMFASAVSKLNGMGSFIYLSKWMKYEFPQAGPCPKAASKGLKLCNWV
  PELDTITCLGLLIDENFTLVLYSLDHLBELKSIARVADGLASESKLSCFVANVVESL
  KLGARN"
  26772..28802
  /gene="F6A14.7"
  /join(26772..27077,27487..27567,27648..28677,28771..28802)
  /gene="F6A14.7"
  /note="Unknown protein; Location of EST G12B1T7,
  gb|W43460"
  /codon_start=1
  /protein_id="AAF27097.1"
  /db_xref="GI:6730702"
  /translation="WVYHVQIRVNGKEVWVVISKIBESDVSLSSIGNAAVTSGIVE
  TQNLKHEVDSDIEQVSEVQPTQSDVASVPDLSSEKIQOEIAAVTVQAAVRYGL
  ENKANSVTHSYLGINKLTGNAPAKLLASSPNVPLSLDLNDSNSIWLNNWSACP
  WKPVQPKASLRKSKQKSPASNPQIVAEAFAPKPKSVKVPSSNLDNSVAQTSSELE
  KPRGFRKYSTSQSVEPFLPSMDNPQVDLEKVRGLKRVHNPVENSIOPLVQIAVE
  KPNGLSESVNAFDEKEDEVAETVQPEELIQHTPLGTNESLDSTLVNQIEESE
  NYMAEKEDVKEETPFKNHKNESAKENQKSKASSVTATQTAEFQESNGNQTSS
  PCIPYMOATKSAAKRLQSGSSPRLGTTEKARRYSLPSSGNSAKITSHSPKTRV
  SNSGKSGNKTETKILLSRENGKATPVEMKR"
  complement(35250..40561)
  /gene="F6A14.8"
  /complement(join(35250..35476,36230..36440,36528..36662,
  36742..36803,36897..37110,37199..37384,37491..37583,
  37674..38073,38169..38296,38844..38987,39095..39217,
  39414..39513,39610..39646,39743..39839,39921..40037,
  40123..40177,40266..40561))
  /gene="F6A14.8"
  /note="Similar to WEB1/SEC31-like protein transport
  protein; Similar to WEB1/SEC31 protein transport protein"
  /codon_start=1
  /protein_id="AAF27099.1"
  /db_xref="GI:6730704"
  /translation="MDCIKSIGRSFAVAIPESPPIAAGTWAGAVDLSFSSSANLEIF

Alignment Scores:
Pred. No.:      9.15e+03      Length:      108767
Score:          38.00         Matches:      8
Percent Similarity: 88.9%      Conservative: 0
Best Local Similarity: 88.9%      Mismatches:  1
Query Match:     86.4%         Indels:      0
DB:              15           Gaps:        0

US-10-774-176-5 (1-9) x AC011809 (1-108767)
Oy 1 PheLeuThrGlyAsnGlnLeuAlaVal 9
|||||
Db 49678 TTTTACAGGGCAATCAATAGCCGTC 49704

RESULT 47
LOCUS      CR762469
DEFINITION Danio rerio chromosome 16 clone CH211-119N15, *** SEQUENCING IN
PROGRESS ***, 10 unordered pieces.
ACCESSION  CR762469
VERSION    CR762469.4 GI:67763654
KEYWORDS   HTG; HTGS_PHASE1.
```

```

SOURCE          Danio rerio (zebrafish)
ORGANISM        Danio rerio
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE       1 (bases 1 to 146219)
AUTHORS         McLay,K.
TITLE           Direct Submission
JOURNAL         Submitted (13-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
                Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                zfsh-help@sanger.ac.uk Clone requests:
                http://www.sanger.ac.uk/Projects/D_rerio/faq.shtmldataeight
                On Jun 14, 2005 this sequence version replaced gi:67624883.
                ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk
                ----- Project Information
Center project name: zC119N15
                ----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 142933 bases at least Q40
Consensus quality: 143582 bases at least Q30
Consensus quality: 144106 bases at least Q20
Insert size: 145319; sum-of-contigs
Quality coverage: 7.39x in Q20 bases; sum-of-contigs Quality
coverage: 7.05x in Q20 bases; agarose-fp
                -----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4466: contig of 4466 bp in length
* 4467 4566: gap of 100 bp
* 4567 17357: contig of 12691 bp in length
* 17358 17357: gap of 100 bp
* 17358 54092: contig of 36735 bp in length
* 54093 54192: gap of 100 bp
* 54193 62657: contig of 8465 bp in length
* 62658 62757: gap of 100 bp
* 62758 82728: contig of 19971 bp in length
* 82729 86417: contig of 3589 bp in length
* 86418 86517: gap of 100 bp
* 86518 105685: contig of 19168 bp in length
* 105686 105785: gap of 100 bp
* 105786 121122: contig of 15337 bp in length
* 121123 121222: gap of 100 bp
* 121223 130003: contig of 8781 bp in length
* 130004 130103: gap of 100 bp
* 130104 146219: contig of 16116 bp in length.

FEATURES             Location/Qualifiers
     source            1..146219
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /chromosome="16"
                     /clone="CH211-119N15"
                     /clone_lib="CHORI-211"
     misc_feature      1..4466
                     /note="assembly fragment:00058
                     fragment chain:1"
     misc_feature      4567..17357
                     /note="assembly fragment:00436
                     fragment chain:1"
     misc_feature      17358..54092
```

/note="assembly_fragment:01315
 fragment_chain:1"
 54193..62657
 /note="assembly_fragment:00126
 fragment_chain:1"
 62758..82728
 /note="assembly_fragment:00611
 fragment_chain:2"
 82829..86417
 /note="assembly_fragment:00029
 fragment_chain:2"
 86518..105685
 /note="assembly_fragment:01067
 fragment_chain:2"
 105786..121122
 /note="assembly_fragment:00217
 fragment_chain:2"
 121223..130003
 /note="assembly_fragment:00325
 fragment_chain:3"
 130104..146219
 /note="assembly_fragment:00830
 fragment_chain:3"

ORIGIN

Alignment Scores: 1.22e+04 Length: 146219
 Pred. No.: 38.00 Matches: 8
 Score: 88.9% Conservative: 0
 Percent Similarity: 88.9% Mismatches: 1
 Best Local Similarity: 86.4% Indels: 0
 Query Match: 14 Gaps: 0
 DB:

US-10-774-176-5 (1-9) x CR762469 (1-146219)

QY 1 PheLeuThrGlyAanGlnLeuAlaVal 9
 |||||
 Db 8410 TTCTTACATCCATCAGTCGCAGTA 8436

RESULT 48
 AC140093/c
 LOCUS AC140093 155001 bp DNA linear HTG 22-MAR-2003
 DEFINITION Bos taurus clone RP42-262C3, WORKING DRAFT SEQUENCE, 8 ordered
 pieces.

ACCESSION AC140093.2 GI:29150354
 VERSION HTG; HTGS PHASE2; HTGS_DRAFT.
 KEYWORDS Bos taurus (cow)
 SOURCE Bos taurus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 155001)
 AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Cariga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
 Haghghi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
 Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.I., Maduro, V.B.,
 Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J.,
 Paquirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
 Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
 Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
 Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative

TITLE Unpublished
 JOURNAL 2 (bases 1 to 155001)
 REFERENCE Green, E.D.

TITLE Direct Submission

TITLE Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
 JOURNAL Grovemont Circle, Gaithersburg, MD 20877, USA

REFERENCE 3 (bases 1 to 155001)
 AUTHORS Green, E.D.
 TITLE Direct Submission

JOURNAL

COMMENT

Submitted (22-MAR-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Mar 22, 2003 this sequence version replaced gi:28460769.

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoo@hgr.nih.gov
 ----- Project Information
 Center project name: dds
 Center clone name: 262C03

The sequence data in this record represents an 'enhanced'
 version of a Phase 2 submission. Specifically, the indicated
 order and orientation of each sequence contig has been
 established using one or more of the following: read-pair
 data from individual subclones, overlaps with neighboring
 clones, alignment with available reference sequence (e.g.,
 human), and/or confirmation by PCR testing. In addition,
 the sequence assembly is based on at least 8X average
 coverage in Q20 bases and has been reviewed to rule out
 gross misassemblies, the low-quality ends of sequence
 contigs have been trimmed away, and each base is associated
 with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 153316 bases at least Q40
 Consensus quality: 153884 bases at least Q30
 Consensus quality: 154212 bases at least Q20
 Insert size: 138000; agarose-fp
 Insert size: 154301; sum-of-contigs
 Quality coverage: 9.57x in Q20 bases; agarose-fp
 Quality coverage: 8.56x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 719: contig of 719 bp in length
 * 720 819: gap of unknown length
 * 14217: contig of 13398 bp in length
 * 14317: gap of unknown length
 * 14318 16197: contig of 1880 bp in length
 * 16198 16297: gap of unknown length
 * 16298 18432: contig of 2135 bp in length
 * 18433 18532: gap of unknown length
 * 18533 32260: contig of 13728 bp in length
 * 32261 32360: gap of unknown length
 * 32361 76744: contig of 4384 bp in length
 * 76744 76844: gap of unknown length
 * 76845 109615: contig of 32771 bp in length
 * 109616 109715: gap of unknown length
 * 109716 155001: contig of 45286 bp in length.

FEATURES

source

1..155001
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-262C3"
 /clone_lib="RP42"

misc_feature

1..76001

/note="clone overlaps with GenBank Accession Number
 AC139631 clone RP42-108E2 (center project name ddr)"

1..719

misc_feature

/note="assembly_fragment
 clone_end:77

vector side:left"

gap 720. .819
/estimated length-unknown

misc_feature 820. .14217
/note="assembly_fragment"

gap 14218. .14317
/estimated length-unknown

misc_feature 14318. .16197
/note="assembly_fragment"

gap 16198. .16297
/estimated length-unknown

misc_feature 16298. .18432
/note="assembly_fragment"

gap 18433. .18532
/estimated length-unknown

misc_feature 18533. .32260
/note="assembly_fragment"

gap 32261. .32360
/estimated length-unknown

misc_feature 32361. .76744
/note="assembly_fragment"

gap 76745. .76844
/estimated length-unknown

misc_feature 76845. .109615
/note="assembly_fragment"

gap 109616. .109715
/estimated length-unknown

misc_feature 109716. .155001
/note="assembly_fragment"

clone end:SP6

vector side:right"

ORIGIN

Alignment Scores:

Pred. No.:	1.29e+04	Length:	155001
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	87.5%	Mismatches:	0
Query Match:	86.4%	Indels:	0
DB:	14	Gaps:	0

US-10-774-176-5 (1-9) x AC140093 (1-155001)

Qy 1 PheLeuThrGlyAsnGlnLeuAla 8
||||:|||||
Db 116046 TTCATTACTGGTAATCAGTTGGCT 116023

RESULT 49
AC158516/c
LOCUS AC158516 167046 bp DNA linear ROD 21-JUN-2005
DEFINITION Mus musculus BAC clone RP24-511A23 from chromosome 9, complete sequence.

ACCESSION AC158516 AC117768
VERSION AC158516.2 GI:63025421
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 167046)
AUTHORS Adams S., Cotton M. and Haglund K.
TITLE The sequence of Mus musculus BAC clone RP24-511A23
JOURNAL Unpublished (2001)

REFERENCE 2 (bases 1 to 167046)
AUTHORS Wilson, R.K.

TITLE Direct Submission
JOURNAL Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 167046)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park

REFERENCE

4 (bases 1 to 167046)
Parkway, St. Louis, MO 63108, USA

WILSON, R.K.
Direct Submission

Submitted (21-JUN-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

On May 4, 2005 this sequence version replaced gi:61656412.

COMMENT

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu>

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0511A23

Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e. phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. The assembly was confirmed by restriction digest.

This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal.

If a sequence meets the criteria of the above statement, it needs no comments or tags. If the criteria are not met, such as ambiguous bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57Bl/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

FEATURES

Location/Qualifiers

1..167046
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/clone="RP24-511A23"
/clone_lib="RP24-511A23"
/note="Sequence derived from PCR product of genomic DNA"
16685..16712
misc_feature
31565..31779
unsure
/note="Unresolved simple sequence repeat."
46721..46808
unsure
/note="Unresolved simple sequence repeat."
142336..142347
unsure
/note="Sequence derived from one plasmid subclone."

ORIGIN

Alignment Scores:	1.38e+04	Length:	167046
Pred. NO.:	38.00	Matches:	7
Score:	88.9%	Conservative:	1

Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 86.4% Indels: 0
 DB: 9 Gaps: 0

US-10-774-176-5 (1-9) x AC158516 (1-167046)

Qy 1 PheLeuThrGlyAaGlnLeuAlaVal 9

Db 110550 TTCTTACGGGACACGATGACCGTG 110524

RESULT 50

AC160108/c

LOCUS

DEFINITION Mus musculus BAC clone RP24-155K11 from chromosome 13, complete

ACCESSION AC160108 AC118718

VERSION AC160108.2 GI:66841676

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 168567)

The sequence of Mus musculus BAC clone RP24-155K11

Unpublished (2001)

2 (bases 1 to 168567)

Wilson, R.K.

Direct Submission

Submitted (19-APR-2005) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 168567)

Wilson, R.K.

Direct Submission

Submitted (01-JUN-2005) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 168567)

Wilson, R.K.

Direct Submission

Submitted (21-JUN-2005) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Jun 1, 2005 this sequence version replaced gi:62734848.

COMMENT

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@watson.wustl.edu

----- Summary Statistics

Center project name: M_BB0155K11

Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e. phred quality
 >30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone, fosmid clone or direct clone walk sequence.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to obtain the consensus sequence. The
 assembly was confirmed by restriction digest.
 This finishing standard has slightly changed from the previous
 Human standard. Specifically, standards for regions of low sequence
 complexity (such as dinucleotide repeats and small unit tandem
 repeats) have been relaxed. These regions are very prevalent in the
 mouse genome, and the return on extended finishing efforts is
 minimal.

If a sequence meets the criteria of the above statement, it needs
 no comments or tags. If the criteria are not met, such as ambiguous
 bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Was Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The BAC Library has been constructed by Pieter de Jong and
 coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
 and/or brain genomic DNA. The clone and detailed information can be
 obtained from Pieter de Jong and coworkers at http://www.chori.org

This sequence is the entire insert of the clone.

FEATURES

source

1..168567
 /location="Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="13"
 /clone_lib="RPCI-24"
 /clone="RP24-155K11"

unseq

19608..19903
 /note="Unresolved simple sequence repeat."

unseq

55213..55321

/note="Sequence derived from one plasmid subclone."

misc_feature

55407..55625

/note="Sequence derived from PCR product of project DNA"

misc_feature

56668..56851

/note="Sequence derived from PCR product of project DNA"

unseq

71549..71707

/note="Unresolved simple sequence repeat."

unseq

139054..139175

/note="Unresolved simple sequence repeat."

misc_feature

149542..149670

/note="Sequence derived from PCR product of project DNA"

unseq

167744..167828

/note="Unresolved simple sequence repeat."

ORIGIN

Alignment Scores:

Pred. No.:	1.39e+04	Length:	168567
Score:	38.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	86.4%	Indels:	0
DB:	9	Gaps:	0

US-10-774-176-5 (1-9) x AC160108 (1-168567)

Qy 1 PheLeuThrGlyAaGlnLeuAlaVal 9

Db 161305 TTTTACAGGAACAAATGCTGTT 161279

Search completed: April 25, 2006, 20:29:24

Job time : 3108.7 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-24
Perfect score: 41
Sequence: 1 AIFLLVLYL 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB spool/US10774176/runat 26052006 091441 24976/app query.fasta 1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40 cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs802h
-USER=US10774176 @CGN 1.1 2389 @runat 26052006 091441 24976 -NCFU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005s:*
- 15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	246	10 ADK11641	Adk11641 Breast ca
2	41	100.0	475	13 ADU11677	Adul1677 Solid tum
3	41	100.0	901	3 AAA27060	Aaa27060 Canine 5T

4	41	100.0	927	6	ABT07721	Abt07721 Breast ca	
5	41	100.0	927	8	ABX76333	Abx76333 Lung canc	
6	41	100.0	927	10	ADB80503	Adb80503 Ovarian c	
7	41	100.0	927	11	ADN38723	Adn38723 Cancer/an	
8	41	100.0	973	8	AAD56198	Aad56198 Human LRR	
9	41	100.0	1156	6	ABV99349	Abv99349 Human NOV	
10	41	100.0	1260	6	ABK87175	Abk87175 cDNA enco	
11	41	100.0	1260	10	ADB97513	Adb97513 Feline 5T	
12	41	100.0	1260	10	ADB97452	Adb97452 DNA encod	
13	41	100.0	1263	3	AAA27058	Aaa27058 Human 5T4	
14	41	100.0	1263	4	AAF89736	Aaf89736 Nucleotid	
15	41	100.0	1263	6	ABK87174	Abk87174 cDNA enco	
16	41	100.0	1281	3	AAA27059	Aaa27059 Mouse 5T4	
17	41	100.0	1331	8	AAD56199	Aad56199 Human LRR	
18	41	100.0	2020	10	ADJ56299	Adj56299 Human CDN	
19	41	100.0	2053	8	ACC51052	Acc51052 Human bla	
20	41	100.0	2053	8	ABX76332	Abx76332 Lung canc	
21	41	100.0	2053	8	AAD56197	Aad56197 Human LRR	
22	41	100.0	2053	8	AAD56200	Aad56200 Human LRR	
23	41	100.0	2053	11	ADN38721	Adn38721 Cancer/an	
24	41	100.0	2053	12	ADL06473	Adl06473 Human tum	
25	41	100.0	2053	12	ADN03961	Adn03961 Antipsoi	
26	41	100.0	2053	13	ADR25444	Adr25444 Breast ca	
27	41	100.0	2053	13	ACN38510	Acn38510 Tumour-as	
28	41	100.0	2053	13	ADV35098	Adv35098 Human CDN	
29	41	100.0	2053	14	AED17761	Aed17761 Fibrotic	
30	41	100.0	2338	5	AAS87175	Aas87175 DNA encod	
31	41	100.0	2359	4	AAK94253	Aak94253 Human ful	
32	41	100.0	2359	12	ADL30831	Adl30831 Full leng	
33	41	100.0	2361	4	AAK94254	Aak94254 Human ful	
34	41	100.0	2361	12	ADI26162	Adi26162 Human CDN	
35	41	100.0	2361	12	ADL30833	Adl30833 Full leng	
36	41	100.0	2557	12	ADI26160	Adi26160 Human CDN	
37	41	100.0	2557	12	ADI26158	Adi26158 Human CDN	
C	38	40	97.6	19142	2	AAX20580	Aax20580 Polynucle
39	37	90.2	204	10	ADH84583	Adh84583 Enterococ	
C	40	37	90.2	1058	4	ABL06215	Ab106215 Drosophoc
41	37	90.2	1421	3	AAC46422	Aac46422 Arabidops	
42	37	90.2	1422	3	AAC36467	Aac36467 Arabidops	
43	37	90.2	1689	13	ADT15316	Adt15316 Plant CDN	
44	37	90.2	3298	4	ABL06214	Ab106214 Drosophoc	
45	37	90.2	6224	6	ABL33308	Ab133308 Human imm	

ALIGNMENTS

RESULT 1
ADK11641
ID ADK11641 standard; DNA; 246 BP.
XX
AC ADK11641;
XX
DT 06-MAY-2004 (first entry)
XX
DE Breast cancer differentially expressed gene product #47.
XX
KW ds; cytostatic; gene therapy; DKFZp5661133 activity inhibitor;
KW Breast cancer; differential expression.
XX
OS Homo sapiens.
XX
FN WO2003057926-A1.
XX
PD 17-JUL-2003.
XX
PF 08-JAN-2003; 2003WO-US000657.
XX
PR 08-JAN-2002; 2002US-0345637P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Hansen R;
XX

DR WPI; 2003-577534/54.
PT Inhibiting a cancerous phenotype of a cell, useful for treating breast
PT cancer comprises contacting a cancerous mammalian cell with an agent for
PT inhibition of DKFZp5661133 activity.
XX
XX
XX Claim 30; SEQ ID NO 47; 257pp; English.
XX
XX The invention relates to a method of inhibiting a cancerous phenotype of
CC a cell comprises contacting a cancerous mammalian cell with an agent for
CC inhibition of DKFZp5661133 activity. The methods are useful for treating
CC cancer e.g. breast cancer. This sequence represents a gene product which
CC is differentially expressed in breast cancer cells. The sequence can be
CC used in the method of the invention.
XX
SQ Sequence 246 BP; 77 A; 49 C; 59 G; 61 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 92.3 Length: 246
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-24 (1-9) x ADU11641 (1-246)
Qy 1 AlAlIlePheLeuLeuValLeuTyrLeu 9
Db 15 GCTATTTTCCTCCTCGTTTGTATTG 41

RESULT 2
ADU11677
ID ADU11677 standard; DNA; 475 BP.
XX
AC ADU11677;
XX
DT 27-JAN-2005 (first entry)
XX
DE Solid tumour prognosis gene seqid 2116.
XX
KW cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
OS Unidentified.
XX
PN WO2004097052-A2.
XX
PD 11-NOV-2004.
XX
XX 29-APR-2004; 2004WO-US013587.
XX
XX 29-APR-2003; 2003US-0466067P.
XX
XX 23-JAN-2004; 2004US-0538246P.
XX
XX (AMHP) WYETH.
XX (STRA/) STRAHS A.
XX
PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dorner AJ;
XX
XX WPI; 2004-804779/79.
XX
XX A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
XX Disclosure; Page; 111pp; English.
XX
XX The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially

CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster profiled by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 183 Length: 475
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-24 (1-9) x ADU11677 (1-475)
Qy 1 AlAlIlePheLeuLeuValLeuTyrLeu 9
Db 381 GCTATTTTCCTCCTCGTTTGTATTG 407

RESULT 3
AAA27060
ID AAA27060 standard; DNA; 901 BP.
XX
AC AAA27060;
XX
DT 22-AUG-2000 (first entry)
XX
DE Canine 5T4 tumour-associated antigen gene.
XX
KW Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
XX Canis sp.
XX
XX Key Location/Qualifiers
FT 1..858
FT /tag= a
FT /product= "5T4 antigen"
FT misc_feature 61..74
FT /tag= b
FT /note= "given in the specification but does not seem to
FT be part of the coding sequence and does not encode any
FT corresponding amino acids"
FT misc_feature 135..146
FT /tag= c

FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 207..216
 /tag= d
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 277..290
 /tag= e
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 351..361
 /tag= f
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 422..436
 /tag= g
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 497..511
 /tag= h
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 572..583
 /tag= i
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 644..653
 /tag= j
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 714..723
 /tag= k
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 784..801
 /tag= l
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 WO200029428-A2.
 25-MAY-2000.
 18-NOV-1999; 99WO-GB003859.
 18-NOV-1998; 98GB-00025303.
 27-JAN-1999; 99GB-00001739.
 30-JUL-1999; 99GB-00017995.
 (OXFO-) OXFORD BIOMEDICA UK LTD.
 Carroll MW, Myers KA;
 WPI; 2000-38735/33.
 P-PSDB; AMY94351.
 Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 response useful in vaccinating against and in treating tumors.
 Disclosure; Page 78-79; 79pp; English.
 The present sequence encodes the canine 5T4 tumour-associated antigen
 (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in

CC carcinomas but has a highly restricted expression pattern in normal adult
 tissues. It appears to be strongly correlated to metastasis in colorectal
 and gastric cancer. 5T4 antigen may therefore be useful in tumour
 diagnosis, targeting and immunotherapy. Mice in which tumours had been
 induced were inoculated with a virus expression vector containing the
 human or murine 5T4 gene sequence. The 5T4 antigen was shown to be
 effective at eliciting an immunotherapeutic anti-tumour response. Both
 the nucleic acid encoding the antigen and the antigen itself can be used
 to elicit an immune response, preferably CTL or an antibody response in a
 subject
 SQ Sequence 901 BP; 178 A; 246 C; 212 G; 153 T; 0 U; 112 Other;
 Alignment Scores:
 Pred. No.: 355 Length: 901
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-24 (1-9) x AAA27060 (1-901)
 QY 1 AlaIlePheLeuLeuValLeuTyrLeu 9
 Db 669 GCCATCTTCTCTACTGTTTGTATTG 695
 RESULT 4
 ABT07721
 ID ABT07721 standard; DNA; 927 BP.
 XX AC ABT07721;
 XX DT 14-NOV-2002 (first entry)
 XX DE Breast cancer-associated gene sequence 29.
 XX Gene; ds; breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.
 XX OS Unidentified.
 XX PN WO200259377-A2.
 XX PD 01-AUG-2002.
 XX PF 24-JAN-2002; 2002WO-US002242.
 XX PR 24-JAN-2001; 2001US-0263965P.
 PR 02-FEB-2001; 2001US-0265928P.
 PR 09-APR-2001; 2001US-00829472.
 PR 09-APR-2001; 2001US-0282698P.
 PR 04-MAY-2001; 2001US-0288590P.
 PR 23-MAY-2001; 2001US-0294443P.
 XX (BOSB-) BOS BIOTECHNOLOGY INC.
 XX PA Mack DH, Gish KC, Afar D;
 XX PI WPI; 2002-583738/62.
 XX DR N-PSDB; ABJ05564.
 XX PT Detecting a breast cancer-associated transcript in a patient's cell,
 PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.
 XX Claim 9; Page 372; 414pp; English.
 XX The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown

CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 366 Length: 927
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-24 (1-9) x ABT07721 (1-927)

Qy 1 AlatlPheLeuValLeuTyrLeu 9
 |||||
 Db 760 GCTATTTTCCTCGTTTGTATTG 786

RESULT 5

ABX76333
 ID ABX76333 standard; DNA; 927 BP.

AC ABX76333;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polynucleotide #197.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified.

PN WO200286443-A2.

XX 31-OCT-2002.

PF 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

DR P-PSDB; ABUS56604.

PT Detecting a lung cancer-associated transcript in a cell from a patient

PT for treating lung cancer, by contacting a biological sample from the

PT patient with a polynucleotide that exhibits increased or decreased

PT expression in lung cancer.

XX Claim 22; Page 336; 453pp; English.

CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 366 Length: 927
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-24 (1-9) x ABX76333 (1-927)

Qy 1 AlatlPheLeuValLeuTyrLeu 9

|||||
 Db 760 GCTATTTTCCTCGTTTGTATTG 786

RESULT 6

ADB80503

ID ADB80503 standard; DNA; 927 BP.

XX ADB80503;

XX 04-DEC-2003 (first entry)

DE Ovarian cancer-associated transcript #34.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;

KW post-operative chemotherapy; radiation therapy; tumour prognosis;

KW pre-cancerous lesion detection; ds; gene.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..927

FT /*tag= a

XX WO2002102235-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

```

DR P-PSDB; ADB80504.
XX
PT Detecting an ovarian cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS Claim 10; Page 297; 332pp; English.
XX
CC The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a
CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 366 Length: 927
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-24 (1-9) x ADB80503 (1-927)
Qy 1 AlallePheLeuValLeuTyLeu 9
Db 760 GCTATTTCTCTCGTTTGTATTG 786

RESULT 7
ADN38723
ID ADN38723 standard; cDNA; 927 BP.
XX
AC ADN38723;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
XX
XX 13-NOV-2001; 2001US-0350666P.
XX
XX 21-NOV-2001; 2001US-0332464P.
XX
XX 29-NOV-2001; 2001US-0334393P.
XX
XX 03-DEC-2001; 2001US-0335394P.
XX
XX 14-DEC-2001; 2001US-0340376P.
XX
XX 08-JAN-2002; 2002US-0347211P.
XX
XX 10-JAN-2002; 2002US-0347349P.
XX
XX 08-FEB-2002; 2002US-0355250P.
XX
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (BOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
DR P-PSDB; ADN38724.
XX
XX Determining the presence or absence of a pathological cell in a patient,
XX useful for diagnosing, prognosing or treating cancer, comprises detecting
XX a nucleic acid in a biological sample.
XX
XX Claim 8; SEQ ID NO 41; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 366 Length: 927
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-24 (1-9) x ADN38723 (1-927)
Qy 1 AlallePheLeuValLeuTyLeu 9
Db 760 GCTATTTCTCTCGTTTGTATTG 786

RESULT 8
AAD56198
ID AAD56198 standard; DNA; 973 BP.
XX
XX AAD56198;
XX
XX 07-AUG-2003 (first entry)
XX
XX Human LRRCAPS related DNA #5.
XX
XX Human; p53 pathway; Leucine rich repeat capricious related protein;
XX LRRCAPS; cancer; gene therapy; ds.
XX
XX Homo sapiens.

```

```

XX PN WO2003035831-A2.
XX PD 01-MAY-2003.
XX PF 21-OCT-2002; 2002WO-US033540.
XX PR 22-OCT-2001; 2001US-0338733P.
XX PR 15-FEB-2002; 2002US-0357600P.
XX PR 01-MAR-2002; 2002US-0361196P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
XX PI Francis-Lang H, Friedman L;
XX DR WPI; 2003-421410/39.
XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
XX PT comprises contacting an assay system comprising a purified leucine rich
XX PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX PS Example 5; Page 74-75; 99pp; English.
XX CC The invention relates to a method of identifying a candidate p53 pathway
XX CC modulating agent. The method involves contacting an assay system
XX CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX CC polypeptide or nucleic acid or its fragment with a test agent and
XX CC detecting a test agent-biased activity, where a difference between the
XX CC test agent-biased activity and the reference activity identifies the test
XX CC agent as a candidate p53 pathway modulating agent. The method is useful
XX CC for identifying a candidate p53 pathway-modulating agent for preparing a
XX CC composition for diagnosing or treating cancer. The invention is useful in
XX CC gene therapy. The present sequence is human LRRCAPS related DNA
XX SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 385 Length: 973
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-24 (1-9) x AAD56198 (1-973)

QY 1 AlailePheLeuValLeuTyrIleu 9
Db 775 GCTATTTTCCTCGTGTGTTTGTATTG 801

RESULT 9
ABV99349
ID ABV99349 standard; DNA; 1156 BP.
XX AC ABV99349;
XX DT 27-JAN-2003 (first entry)
XX DE Human NOV8a coding sequence.
XX KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
KW neotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX OS Homo sapiens.

```

(CURA-) CURAGEN CORP.

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 PI Pena CE, Burges CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 PI Zhong M;
 DR WPI; 2002-732824/79.
 DR P-PSDB; ABP70071.
 XX
 XX New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 PT disorders, and asthma.
 XX
 PS Claim 16; Page 114-115; 619pp; English.
 XX
 XX The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, hematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods
 XX
 SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 460 Length: 1156
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-24 (1-9) x ABV99349 (1-1156)
 Qy 1 AlallePheLeuValLeuTyrLeu 9
 Db 991 GCTATTTCTCTCGTTTGTATTG 1017
 RESULT 10
 ABK87175
 ID ABK87175 standard; cDNA; 1260 BP.
 XX
 AC ABK87175;
 XX
 XX 07-OCT-2002 (first entry)
 XX
 XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
 DE
 XX Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX
 OS Felis sp.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a

FT
 XX /product= "5T4 protein"
 PN WO200238612-A2.
 XX
 PD 16-MAY-2002.
 XX
 XX 13-NOV-2001; 2001WO-GB005004.
 PF
 XX 13-NOV-2000; 2000WO-GB004317.
 PR
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 PA
 XX Myers K, Drury N, Carroll M;
 PI
 XX WPI; 2002-557449/59.
 DR P-PSDB; AAU98694.
 DR
 XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.
 PS
 XX Claim 4; Page 68; 68pp; English.
 XX
 XX The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes feline 5T4 protein
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 504 Length: 1260
 Score: 41.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-24 (1-9) x ABK87175 (1-1260)
 Qy 1 AlallePheLeuValLeuTyrLeu 9
 Db 1099 GCCATTTCTTACTGGTTTGTACTTG 1125
 RESULT 11
 ADB97513
 ID ADB97513 standard; DNA; 1260 BP.
 XX
 XX ADB97513;
 XX
 XX 04-DEC-2003 (first entry)
 DT
 XX Feline 5T4 antigen DNA.
 DE
 XX Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.
 XX
 OS Unidentified.
 XX


```
KW ds.
XX Homo sapiens.
XX WO200029428-A2.
XX PD 25-MAY-2000.
XX PF 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MW, Myers KA;
XX DR WPI; 2000-387735/33.
XX PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the human 5T4 tumour-associated antigen
XX (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The 5T4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject
XX SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 505 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-24 (1-9) x AAA27058 (1-1263)
Qy 1 AlaIlePheLeuLeuValLeuTyrLeu 9
Db 1102 GCTATTTTCCTCGTGGTTTGTATTG 1128

RESULT 14
ID AAF89736 standard; DNA; 1263 BP.
XX AC AAF89736;
XX DT 23-JUL-2001 (first entry)
XX DE Nucleotide sequence of canine 5T4 protein.
XX KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
XX KW hypersensitivity; autoimmune disease; central nervous system disorder;
XX KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
XX KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
XX KW Helicobacter-related disease; immune disorder; ss.
XX OS Canis sp.
XX Key Location/Qualifiers
XX CDS 1..1263
```

```
FT /*tag= a
XX /product= "5T4"
XX PN WO200136486-A2.
XX PD 25-MAY-2001.
XX PF 13-NOV-2000; 2000WO-GB004317.
XX PR 18-NOV-1999; 99WO-GB003859.
XX PR 15-FEB-2000; 2000GB-00003527.
XX PR 02-MAR-2000; 2000GB-00005071.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
XX PI Myers KA;
XX DR WPI; 2001-343805/36.
XX DR P-PSDB; AAB83839.
XX PT Use of single chain antibody capable of recognizing a disease associated
XX molecule for manufacturing a medicament for preventing and/or treating a
XX disease condition associated with disease associated molecule.
XX PS Disclosure; Fig 26; 118pp; English.
XX CC The specification describes the use of a single chain antibody (ScFv),
XX CC which is capable of recognizing a disease associated molecule in the
XX CC manufacture of a medicament for the prevention and treatment of a disease
XX CC condition. The ScFv antibody is useful in the manufacture of a
XX CC medicament, for affecting a disease in vivo, for preparing a
XX CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
XX CC treatment of a disease. The ScFv antibody is also useful for treating
XX CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
XX CC diseases, cancers, central nervous system disorders including Parkinson's
XX CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
XX CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
XX CC related diseases, and other immune disorders. The present sequence
XX CC encodes a 5T4 protein, which is used to produce ScFv of the invention
XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 505 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-24 (1-9) x AAF89736 (1-1263)
Qy 1 AlaIlePheLeuLeuValLeuTyrLeu 9
Db 1102 GCCATCTTCCTACTGTTTGTATTG 1128

RESULT 15
ID AAK87174 standard; cDNA; 1263 BP.
XX AC AAK87174;
XX DT 07-OCT-2002 (first entry)
XX DE cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
XX KW Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
XX KW cell proliferative disorder; infection; inflammatory condition;
XX KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
XX KW foetal abnormality; foetal sex determination; gene; ss.
XX OS Canis sp.
```


GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds
(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-24

Perfect score: 41

Sequence: 1 AIFLLVLYL 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB spool/US10774176/runat_26052006_091443_24987/app_query.fasta_1
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORES=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs04
-USER=US10774176 @CGN 1 1 7524 @runat_26052006_091443_24987 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *

1: gb env: *

2: gb pat: *

3: gb ph: *

4: gb pl: *

5: gb pr: *

6: gb ro: *

7: gb sts: *

8: gb sy: *

9: gb un: *

10: gb vl: *

11: gb ov: *

12: gb htg: *

13: gb in: *

14: gb on: *

15: gb ba: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	290	2	CQ687716	Sequence
2	41	100.0	475	2	CQ920916	Sequence
3	41	100.0	901	2	BD249733	Polypepti

4	41	100.0	901	2	AX025013	Sequence
5	41	100.0	901	2	AX316088	Sequence
6	41	100.0	927	2	AX829164	Sequence
7	41	100.0	1156	2	DD161112	Novel Ant
8	41	100.0	1260	2	AX467373	Sequence
9	41	100.0	1260	2	AX821533	Sequence
10	41	100.0	1260	2	AX821548	Sequence
11	41	100.0	1263	2	BD249731	Polypepti
12	41	100.0	1263	2	AX025011	Sequence
13	41	100.0	1263	2	AX149553	Sequence
14	41	100.0	1263	2	AX316086	Sequence
15	41	100.0	1263	2	AX467371	Sequence
16	41	100.0	1281	2	BD249732	Polypepti
17	41	100.0	1281	2	AX025012	Sequence
18	41	100.0	1281	2	AX316087	Sequence
19	41	100.0	2053	2	CQ731678	Sequence
20	41	100.0	2053	2	DD174290	METHODS O
21	41	100.0	2053	5	HS5740A	Homo sapien
22	41	100.0	2333	6	AF063939	Rattus no
23	41	100.0	2359	2	BD127282	Primer fo
24	41	100.0	2359	2	CQ782724	Sequence
25	41	100.0	2359	5	AK074786	Homo sapi
26	41	100.0	2361	2	AX961916	Sequence
27	41	100.0	2361	2	BD127283	Primer fo
28	41	100.0	2361	2	CQ782726	Sequence
29	41	100.0	2361	5	AK074790	Homo sapi
30	41	100.0	2361	6	BC087011	Rattus no
31	41	100.0	2379	5	BC037161	Homo sapi
32	41	100.0	2423	6	BC058198	Mus muscu
33	41	100.0	2557	2	AX961912	Sequence
34	41	100.0	2557	2	AX961914	Sequence
35	41	100.0	2714	5	AB168308	Macaca fa
36	41	100.0	5551	5	HSA012159	Homo sapi
37	41	100.0	7942	6	MMU012160	Homo sapi
38	41	100.0	110000	4	AP008208	Continuation (358
39	41	100.0	120743	4	AP004082	Oryza sat
40	41	100.0	121909	5	HSJ492P14	Human DNA
41	41	100.0	129010	5	AL589666	Human DNA
42	41	100.0	167046	6	AC158516	Mus muscu
43	41	100.0	208985	12	AC155741	Bos tauru
44	41	100.0	210237	12	AC128294	Rattus no
45	41	100.0	218269	12	AC156733	Bos tauru

ALIGNMENTS

RESULT 1	CQ687716	Sequence 32642 from Patent WO02070737.	290 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ687716	Sequence 32642 from Patent WO02070737.				
DEFINITION	CQ687716	Sequence 32642 from Patent WO02070737.				
ACCESSION	CQ687716	Sequence 32642 from Patent WO02070737.				
VERSION	CQ687716.1	GI:42218962				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1.	Liew,C.C., Marshall,W.E. and Zhang,H.				
AUTHORS		Compositions and methods relating to osteoarthritis				
TITLE		Patent: WO 02070737-A 32642 12-SEP-2002;				
JOURNAL		Chondrogene Inc. (CA)				
FEATURES		Location/Qualifiers				
source		1..290				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				

ORIGIN

Alignment Scores:	8.65	Length:	290
Pred. No.:	41.00	Matches:	9
Score:			


```

DB:
US-10-774-176-24 (1-9) x AX025013 (1-901)
Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 669 GCCATCTTCCTACTGCTTTGTATTG 695

RESULT 5
AX316088
LOCUS AX316088 901 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 3 from Patent EP1160323.
ACCESSION AX316088
VERSION AX316088.1 GI:17899280
KEYWORDS
SOURCE
ORGANISM
Canis sp.
Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 3 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..901
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"
ORIGIN
Alignment Scores:
Pred. No.: 21.2 Length: 901
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-24 (1-9) x AX316088 (1-901)
Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 669 GCCATCTTCCTACTGCTTTGTATTG 695

RESULT 6
AX829164
LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Mack,D.H., Gish,K.C. and Afar,D.
TITLE Methods of diagnosis of breast cancer, compositions and methods of
screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)
FEATURES
source
1..927
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 21.7 Length: 927

```

```

Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-24 (1-9) x AX829164 (1-927)
Qy 1 AlallePheLeuValLeuTyrLeu 9
Db 760 GCTATTTTCCTCTGCTTTGTATTG 786

RESULT 7
DD161112
LOCUS DD161112 1156 bp DNA linear PAT 23-NOV-2005
DEFINITION Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
Encoding The Antigens, and Methods of Use.
ACCESSION DD161112
VERSION DD161112.1 GI:83967439
KEYWORDS JP 2005508604-A/23.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Padigar,M., Shenoy,S.G., Pochart,P.F., Kekuda,R., Gusev,V.Y.,
Zhong,M., Jr,R.J.T., Casman,S.J., Li,L., Miller,C.E.,
Patturajan,M., Anderson,D.W., Malyankar,U.M., Voss,E.Z.,
Spaderna,S.K., Gorman,L., Spytek,K.A., Liu,X., Burgess,C.E.,
Pena,C.E.A., Gerlach,V., Smithson,G., Mezes,P.D., Rastelli,L.,
Boldog,F.L., Guo,X., Vernet,C.A.M., Gangolli,E.A., Tchernev,V.T.
and Zerhusen,B.D.
TITLE Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
Encoding The Antigens, and Methods of Use
JOURNAL Patent: JP 2005508604-A 23 07-APR-2005;
Muralidhara Padigar,Suresh Shenoy,Ramesh Kekuda Vladimir Gusev,
Pascale Pochart,Mei Zhong,Luca Rastelli,Peter Mezes, Glennda
Smithson, Xiaojia Guo,Valerie Gerlach,Stacie Casman, Ferenc
Boldog,Li Li, Bryan Zerhusen,Velizar Tchernev,Esha Gangolli, Corine
Vernet, Carol Pena, Catherine Burgess,Xiaohong Liu,Kimberly
Spytek,Linda Gorman, Steven Spaderna,Edward Voss,Uriel
Malyankar,David Anderson, Meera Patturajan,Charles Miller,Raymond J
Taupier Jr
COMMENT
OS Homo sapiens
PN JP 2005508604-A/23
PD 07-APR-2005
PF 08-MAR-2002 JP 2002571827
PR 19-JUN-2001 US 60/299310,18-JUN-2001 US 60/299027, PR
31-MAY-2001 US 60/294889,31-MAY-2001 US 60/294899, PR
30-MAY-2001 US 60/294485,09-MAR-2001 US 60/274849, PR
13-MAR-2001 US 60/275579,03-MAY-2001 US 60/288342, PR
02-MAY-2001 US 60/288066,30-APR-2001 US 60/287424, PR
13-APR-2001 US 60/283675,04-APR-2001 US 60/281194, PR
02-APR-2001 US 60/280802,02-APR-2001 US 60/280822, PR
03-MAY-2001 US 60/288528,15-MAY-2001 US 60/291190, PR
16-MAY-2001 US 60/291240,16-MAY-2001 US 60/291099, PR
12-MAR-2001 US 60/275235,08-MAR-2001 US 60/274101, PR
08-MAR-2001 US 60/274281,08-MAR-2001 US 60/274322, PR
08-MAR-2001 US 60/274194,02-APR-2001 US 60/280900, PR
30-MAR-2001 US 60/280233,30-MAR-2001 US 60/279995, PR
20-MAR-2001 US 60/277327,20-MAR-2001 US 60/277338, PR
19-MAR-2001 US 60/276994,16-MAR-2001 US 60/276776, PR
14-MAR-2001 US 60/276000,13-MAR-2001 US 60/275601, PR
13-MAR-2001 US 60/275578,20-MAR-2001 US 60/277239, PR
20-MAR-2001 US 60/277321,21-MAR-2001 US 60/277791, PR
22-MAR-2001 US 60/277833,23-MAR-2001 US 60/278152, PR
26-MAR-2001 US 60/278894,27-MAR-2001 US 60/278999, PR
27-MAR-2001 US 60/279036,28-MAR-2001 US 60/279344, PR
19-JUN-2001 US 60/299303,10-JUL-2001 US 60/304354, PR
31-JUL-2001 US 60/309198,03-DEC-2001 US 60/337426, PR
03-DEC-2001 US 60/338092,21-NOV-2001 US 60/332094, PR
14-NOV-2001 US 60/333272,14-NOV-2001 US 60/332271, PR

```

```

14-NOV-2001 US 60/332272,14-NOV-2001 US 60/332172, PR
14-NOV-2001 US 60/333184,31-OCT-2001 US 60/335301, PR
18-OCT-2001 US 60/330380,27-SEP-2001 US 60/325430, PR
27-SEP-2001 US 60/325681,12-SEP-2001 US 60/318770, PR
10-SEP-2001 US 60/318462,03-JAN-2002 US 60/345705, PR
04-DEC-2001 US 60/337185,08-MAR-2002 US 10/093463, PR
16-AUG-2001 US 60/312903
PI muralidhara padigar,suresh g shenoy,pascale f-g pochart, PI
remesh kekuda,
PI vladimir y gusev,mei zhong,raymond j taupier jr,stacie j PI
casman,li li,
PI charles e miller,meera patturajan,david w anderson,uriel m PI
PI malyankar,
PI edward z voss,steven k spaderna,linda gorman,kimberly PI a
spytek,
PI xiaohong liu,catherine e burgess,carol e a pena,valerie PI
gerlach,
PI glenda smithson,peter d mezes,luca rastelli,ferenc l boldog,
PI xiaojia guo,
PI corine a m vernet,esha a gangolli,velizar t tchernev,bryan d
PI zerhusen
CC
FH Key Location/Qualifiers
FT CDS (1)..(1150).
FEATURES
source
1..1156
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 25.9 Length: 1156
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-24 (1-9) x DD161112 (1-1156)
QY 1 AlaiPheLeuValLeuTyrlieu 9
Db 991 GCTATTTTCCTCTGCTGTTTGTATTG 1017
RESULT 8
AX467373 LOCUS 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
Location/Qualifiers
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"
ORIGIN
Alignment Scores:
Pred. No.: 27.7 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0

```

```

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-24 (1-9) x AX467373 (1-1260)
QY 1 AlaiPheLeuValLeuTyrlieu 9
Db 1099 GCCATTTTCTTACTGTTTGTACTTG 1125
RESULT 9
AX821533 LOCUS 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
TITLE MHC class I peptide epitopes from the human St4 tumor-associated
antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
Location/Qualifiers
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 27.7 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-24 (1-9) x AX821533 (1-1260)
QY 1 AlaiPheLeuValLeuTyrlieu 9
Db 1099 GCCATTTTCTTACTGTTTGTACTTG 1125
RESULT 10
AX821548 LOCUS 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
TITLE MHC class II peptide epitope of St4 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
Location/Qualifiers
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN

```

Alignment Scores:
Pred. No.: 27.7 Length: 1260
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-24 (1-9) x AX821548 (1-1260)

Qy 1 AlallePheLeuLeuValLeuTyrLeu 9
|||||
Db 1099 GCCATTTCCTCTACTGGTTTGTATTG 1125

RESULT 11
BD249731 LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD BIOMEDICA LTD (GB)

FEATURES
source 1. .1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 27.8 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-24 (1-9) x AX025011 (1-1263)

Qy 1 AlallePheLeuLeuValLeuTyrLeu 9
|||||
Db 1102 GCTATTTTCCTCTCTGGTTTGTATTG 1128

RESULT 13
AX149553 LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
AUTHORS Kingsman, A.O., Kingsman, S.M., Bebbington, C.R., Carroll, M.W., Ellard, F.M. and Myers, K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source 1. .1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="574"

ORIGIN
Alignment Scores:
Pred. No.: 27.8 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-24 (1-9) x AX149553 (1-1263)

Qy 1 AlallePheLeuLeuValLeuTyrLeu 9
|||||
Db 1102 GCCATCTTCCTCTCTGGTTTGTATTG 1128

RESULT 14
AX316086 LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001

Search completed: May 27, 2006, 19:35:25
Job time : 3359.6 secs

```
DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE St4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1263
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 27.8 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-24 (1-9) x AX316086 (1-1263)
Qy 1 AlailePheLeuValLeuTyrLeu 9
Db 1102 GCTATTTCCTCCTGGTTTGTATTG 1128
RESULT 15
AX467371
LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS
SOURCE Canis sp.
ORGANISM Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1263
/mol_type="unassigned DNA"
/db_xref="taxon:9616"
ORIGIN
Alignment Scores:
Pred. No.: 27.8 Length: 1263
Score: 41.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-24 (1-9) x AX467371 (1-1263)
Qy 1 AlailePheLeuValLeuTyrLeu 9
Db 1102 GCCATCTTCCTACTCGTTTGTATTG 1128
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-23

Perfect score: 43

Sequence: 1 SLQTSYVFL 9

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 524920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB spool/US10774176/runat 26052006.091441.24976/app query.fasta 1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0-
-UNIT5-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10774176 @CGN 1.1 2389 @runat 26052006.091441.24976 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005as:*
- 15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	43	100.0	65	6	ABN56274	Mouse spl
2	43	100.0	108	10	ACD97670	Human col
3	43	100.0	453	5	AAS87174	Human DNA encod

4	43	100.0	475	13	ADU11677	Adult1677 Solid tum
5	43	100.0	901	3	AAA27060	Canine 5T
6	43	100.0	927	6	ABT07721	Breast ca
7	43	100.0	927	8	ABX76333	Lung canc
8	43	100.0	927	10	ADB80503	Ovarian c
9	43	100.0	927	11	ADN38723	Cancer/an
10	43	100.0	973	8	AA56198	Human LRR
11	43	100.0	1156	6	ABV99349	Human NOV
12	43	100.0	1260	6	ABK87175	CDNA enco
13	43	100.0	1260	10	ADB97513	Feline 5T
14	43	100.0	1260	10	ADB97452	DNA encod
15	43	100.0	1263	3	AAA27058	Human 5T4
16	43	100.0	1263	4	AAF89736	Nucleotid
17	43	100.0	1263	6	ABK87174	CDNA enco
18	43	100.0	1281	3	AAA27059	Mouse 5T4
19	43	100.0	1331	8	AA56199	Human LRR
20	43	100.0	2020	10	ADJ56299	Human CDN
21	43	100.0	2053	8	ACC51052	Human bla
22	43	100.0	2053	8	ABX76332	Lung canc
23	43	100.0	2053	8	AA56197	Human LRR
24	43	100.0	2053	8	AA56200	Human LRR
25	43	100.0	2053	11	ADN38721	Cancer/an
26	43	100.0	2053	12	ADL06473	Human tum
27	43	100.0	2053	12	ADN03961	Antipgori
28	43	100.0	2053	13	ADR25444	Breast ca
29	43	100.0	2053	13	ACN38510	Tumour-as
30	43	100.0	2053	13	ADV35098	Human CDN
31	43	100.0	2053	14	AED17761	Fibrotic
32	43	100.0	2338	5	AAS87175	DNA encod
33	43	100.0	2359	4	AAK94253	Human ful
34	43	100.0	2359	12	ADL30831	Full leng
35	43	100.0	2361	4	AAK94254	Human ful
36	43	100.0	2361	12	ADL26162	Human CDN
37	43	100.0	2361	12	ADL30833	Full leng
38	43	100.0	2557	12	ADI26160	Human CDN
39	43	100.0	2557	12	ADI26158	Human CDN
40	39	90.7	14952	10	ACF69283	Phototrab
41	39	90.7	100779	10	ACF65386	Continuation (7 of
42	39	90.7	110000	10	ACF67367	Continuation (23 o
43	38	88.4	117730	14	AD212550	Human can
44	37	86.0	565	13	ADX45743	Plant ful
45	37	86.0	2848	2	AAT75704	Murine le

ALIGNMENTS

RESULT 1
ABN56274
ID ABN56274 standard; DNA; 65 BP.
AC
XX
AC ABN56274;
XX
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:29022.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB001903.
XX
PR 28-JUL-2000; 2000US-0221607P.
XX
PR 02-MAY-2001; 2001US-0287724P.
XX
PA (COMP-) COMPUGEN INC.
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which

PT selectively hybridize to mRNAs transcribed from a transcription unit of a

PT genome, useful for detecting tissue-, pathology-, and developmental-

PT specific genes.

XX Example 1; SEQ ID NO 29022; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting

CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-

CC)transcriptome comprises messenger RNAs transcribed from multiple

CC transcription units that populate a genome. The library comprises several

CC oligonucleotides, each capable of hybridising selectively to a set of

CC messenger RNAs transcribed from a given transcription unit of the genome,

CC which encodes one or more messenger RNA splice variants. The

CC oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or

CC quantitatively characterising the corresponding transcriptome, and in

CC detecting RNA transcripts and splice variants of human or animal

CC transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a particular

CC biological or pathological state, and so allowing the detection of tissue

CC - and pathology-specific genes such as those genes only expressed in

CC specific tissue under a specific pathological condition; to detect

CC developmental specific genes; and to detect RNA transcripts and splice

CC variants of a transcriptome of a patient suffering from a particular

CC disorder. ABN27453 to ABN59589 represent oligonucleotide sequences from

CC rats, humans and mice, which are used in the exemplification of the

CC present invention. N.B. The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 65 BP; 9 A; 21 C; 12 G; 23 T; 0 U; 0 Other;

SQ Sequence 65 BP; 9 A; 21 C; 12 G; 23 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0.822	Length:	65
Score:	43.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-23 (1-9) x ABN56274 (1-65)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 29 TCCCTGCAGACTTCCTATGCTCTCCTA 55

RESULT 2

ACD97670

ID ACD97670 standard; cDNA; 108 BP.

XX ACD97670;

XX 23-SEP-2003 (first entry)

XX Human colon cancer cell expressed cDNA #6082.

DE Open reading frame detection; genome sequencing; colon cancer;

KW breast cancer; population genome analysis; genetic shift; cancer;

KW antibiotic resistance; antibiotic non-tolerance; congenital disease;

KW agriculture; food crop genome; resistance gene; retrovirus;

KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;

KW gene; sa.

XX Homo sapiens.

OS

XX US2002155438-A1.

PN

XX 24-OCT-2002.

PD

XX

PF 27-SEP-1999; 99US-00406117.

XX

PR 20-NOV-1998; 98US-00196716.

XX

PA (SIMP/) SIMPSON A J G.

PA (NETO/) NETO E D.

PA (BREN/) BRENTANI R R.

XX

PI Simpson AJG, Neto ED, Brentani RR;

XX WPI; 2003-182626/18.

DR

XX Determining open reading frames of genome of an organism e.g. a human

PT suffering from cancer involves use of single oligonucleotide primer at

PT low stringency for preparing single-stranded cDNA from mRNA of

PT individual.

XX Example 9; Page 866; 959pp; English.

PS The invention describes a method of determining open reading frames in

XX the genome of organism, comprising contacting mRNA from cell of organism

CC with a single oligonucleotide primer (I) at low stringency, preparing

CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying

CC cDNA, sequencing the product, and repeating the contacting, preparing

CC and amplifying steps with different primers and sequencing, resulting

CC nucleic acids. The method is useful for: determining that a known

CC nucleotide sequence from a genome of an organism corresponds to a

CC nucleotide sequence of an open reading frame; for preparing a contig.

CC nucleic acid molecule from a genome of an organism; and for sequencing

CC all or part of a genome of an organism. mRNA is obtained from mammalian

CC or human cell which is associated with a pathological condition e.g. a

CC colon cancer or breast cancer cell. The method is useful for analyses of

CC populations of subjects and can be used to carry out genetic analyses of

CC large or small populations. further, it can be used to study living

CC systems to determine if, e.g. there have been genetic shifts which render

CC an individual or population more or less likely to be afflicted with

CC diseases such as cancer, to determine antibiotic resistance or non-

CC tolerance, and so forth. The method can also be used in the study of

CC congenital diseases, and the risk of affliction to a foetus, as well as

CC the study of whether the conditions are likely to be passed to offspring

CC through ova or sperm. The analyses for pathological conditions can be

CC carried out in all animals, plants, birds, fish, etc. Using this method,

CC in the area of agriculture, for example the genomes of food crops can be

CC studied to determine if resistance genes are present, defects in plant

CC genomes can also be studied in this way. Similarly, the method permits

CC determination of the pathogens which integrate into the genome, such as

CC retroviruses and other integrating viruses such as influenza virus, have

CC undergone shifts or mutations, which may require different approaches to

CC therapy. This method is also applied to eukaryotic pathogens, such as

CC trypanosomes, different types of Plasmodium, etc. The method essentially

CC eliminates sequencing of non-coding portions. This sequence represents a

CC polynucleotide isolated from human colon cancer cell cDNA library

XX

SQ Sequence 108 BP; 18 A; 33 C; 22 G; 35 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.46	Length:	108
Score:	43.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-774-176-23 (1-9) x ACD97670 (1-108)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 60 TCCCTGCAGAACTCTATGCTCTCCTG 86

RESULT 3

AAS87174

ID AAS87174 standard; cDNA; 453 BP.

XX

AC AAS87174;
XX
XX
DT 13-FEB-2002 (first entry)
XX
XX
DE DNA encoding novel human diagnostic protein #22978.
XX
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200175067-A2.
XX
XX
PD 11-OCT-2001.
XX
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
PI Drmanac RT, Liu C, Tang VT;
XX
XX
DR WPI; 2001-639362/73.
DR P-FSDB; ABG22987.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX
PS Claim 1; SEQ ID NO 22978; 103pp; English.
XX
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 453 BP; 108 A; 111 C; 113 G; 121 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.34 Length: 453
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-23 (1-9) x AAS87174 (1-453)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 238 TCCCTGCAACCTCTTATGCTCTTCCTG 264

RESULT 4

ADU11677
ID ADU11677 standard; DNA; 475 BP.
XX
XX
AC ADU11677;
XX
XX
DT 27-JAN-2005 (first entry)
XX
XX
DE Solid tumour prognosis gene seqid 2116.
XX
XX
KW cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
XX
OS Unidentified.
XX
XX
PN WO2004097052-A2.
XX
XX
PD 11-NOV-2004.
XX
XX
PF 29-APR-2004; 2004WO-US013587.
XX
XX
PR 29-APR-2003; 2003US-0466067P.
PR 23-JAN-2004; 2004US-0538246P.
XX
XX
PA (AMHP) WYETH.
PA (STRA/) STRAHS A.
XX
XX
PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dörner AJ;
XX
XX
DR WPI; 2004-804779/79.
XX
XX
PT A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
XX
PS Disclosure; Page; 111pp; English.
XX
XX
CC The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.: 7.74 Length: 475
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-23 (1-9) x ADU11677 (1-475)
QY 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 330 TCCCTGCAACCTCTTAAGTCTTCCTG 356

RESULT 5
AAA27060
ID AAA27060 standard; DNA; 901 BP.
XX AAA27060;
XX 22-AUG-2000 (first entry)
XX Canine 5T4 tumour-associated antigen gene.
XX Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX ds.
XX Canis sp.
XX
XX Key Location/Qualifiers
XX CDS 1..858
XX /tag= a
XX /product= "5T4 antigen"
XX misc_feature 61..74
XX /tag= b
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 135..146
XX /tag= c
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 207..216
XX /tag= d
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 277..290
XX /tag= e
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 351..361
XX /tag= f
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 422..436
XX /tag= g
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 497..511
XX /tag= h
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 572..583
XX /tag= i
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
```

```
FT corresponding amino acids"
FT 644..653
FT /tag= j
FT /note= "given in the specification but does not seem to
FT be part of the coding sequence and does not encode any
FT corresponding amino acids"
FT 714..723
FT /tag= k
FT /note= "given in the specification but does not seem to
FT be part of the coding sequence and does not encode any
FT corresponding amino acids"
FT 784..801
FT /tag= l
FT /note= "given in the specification but does not seem to
FT be part of the coding sequence and does not encode any
FT corresponding amino acids"
XX WO200029428-A2.
XX 25-MAY-2000.
XX 18-NOV-1999; 99WO-GB003859.
XX 18-NOV-1998; 98GB-00025303.
XX 27-JAN-1999; 98GB-00001739.
XX 30-JUL-1999; 99GB-00017995.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX Carroll MW, Myers KA;
XX WPI; 2000-387735/33.
XX P-PSDB; AAY94351.
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX response useful in vaccinating against and in treating tumors.
XX Disclosure; Page 78-79; 79pp; English.
XX The present sequence encodes the canine 5T4 tumour-associated antigen
XX (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX carcinomas but has a highly restricted expression pattern in normal adult
XX tissues. It appears to be strongly correlated to metastasis in colorectal
XX and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX induced were inoculated with a virus expression vector containing the
XX human or murine 5T4 gene sequence. The 5T4 antigen was shown to be
XX effective at eliciting an immunotherapeutic anti-tumour response. Both
XX the nucleic acid encoding the antigen and the antigen itself can be used
XX to elicit an immune response, preferably CTL or an antibody response in a
XX subject
XX SQ Sequence 901 BP; 178 A; 246 C; 212 G; 153 T; 0 U; 112 Other;

Alignment Scores:
Pred. No.: 15.9 Length: 901
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-23 (1-9) x AAA27060 (1-901)
QY 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 608 TCCCTGCAACCTCTTAAGTCTTCCTA 634

RESULT 6
ABT07721
ID ABT07721 standard; DNA; 927 BP.
XX
XX AC ABT07721;
```

XX 14-NOV-2002 (first entry)
 XX
 XX
 DE Breast cancer-associated gene sequence 29.
 XX
 KW Gene; ds; breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.
 XX
 OS Unidentified.
 XX WO200259377-A2.
 PN
 XX
 PD 01-AUG-2002.
 XX
 XX 24-JAN-2002; 2002WO-US002242.
 PF
 XX 24-JAN-2001; 2001US-0263965P.
 PR
 XX 02-FEB-2001; 2001US-0265928P.
 PR
 XX 09-APR-2001; 2001US-00829472.
 PR
 XX 09-APR-2001; 2001US-0282698P.
 PR
 XX 04-MAY-2001; 2001US-0288590P.
 PR
 XX 29-MAY-2001; 2001US-0294443P.
 PR
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Mack DH, Gish KC, Afar D;
 PI
 XX WPI; 2002-583738/62.
 DR
 XX N-PSDB; ABJ05564.
 DR
 XX
 XX Detecting a breast cancer-associated transcript in a patient's cell,
 PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.
 XX
 XX Claim 9; Page 372; 414pp; English.
 PS
 XX
 XX The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
 CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 16.5 Length: 927
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-23 (1-9) x ABT07721 (1-927)
 Qy 1 SerLeuGlnThrSerTyrrvalPheLeu 9
 Db 709 TCCTGCAACCTCTTATGCTCTCTG 735
 RESULT 7
 ABX76333
 ID ABX76333 standard; DNA; 927 BP.
 XX

AC ABX76333;
 XX
 XX 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polynucleotide #197.
 XX
 KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 XX WO200286443-A2.
 PN
 XX 31-OCT-2002.
 PD
 XX 18-APR-2002; 2002WO-US012476.
 PF
 XX 18-APR-2001; 2001US-0284770P.
 PR
 XX 10-MAY-2001; 2001US-0290492P.
 PR
 XX 09-NOV-2001; 2001US-0339245P.
 PR
 XX 13-NOV-2001; 2001US-0350666P.
 PR
 XX 29-NOV-2001; 2001US-0334370P.
 PR
 XX 12-APR-2002; 2002US-0372246P.
 PR
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 XX Aziz N, Murray R;
 PI
 XX WPI; 2003-093161/08.
 DR
 XX P-PSDB; ABU56604.
 DR
 XX
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 XX Claim 22; Page 336; 453pp; English.
 PS
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, and
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 16.5 Length: 927
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-23 (1-9) x ABX76333 (1-927)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 Db 709 TCCTGCAACCTCTTATGCTTCCTG 735
 RESULT 8
 ADB80503
 ID ADB80503 standard; DNA; 927 BP.
 AC ADB80503;
 XX
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Ovarian cancer-associated transcript #34.
 XX
 KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; ds; gene.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..927
 FT /*tag= a
 XX
 XX WO2002102235-A2.
 XX
 XX 27-DEC-2002.
 XX
 XX 18-JUN-2002; 2002WO-US019297.
 XX
 XX 18-JUN-2001; 2001US-0299234P.
 XX 27-AUG-2001; 2001US-0315287P.
 XX 05-SEP-2001; 2001US-0317544P.
 XX 13-NOV-2001; 2001US-0350666P.
 XX 12-APR-2002; 2002US-0372246P.
 XX
 XX (E0SB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Mack DH, Gish KC;
 XX
 XX WPI; 2003-167431/16.
 DR P-PSDB; ADB80504.
 XX
 PT Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.
 XX
 PS Claim 10; Page 297; 332pp; English.
 XX
 CC The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule.
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 16.5 Length: 927
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-23 (1-9) x ADB80503 (1-927)
 Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 Db 709 TCCTGCAACCTCTTATGCTTCCTG 735
 RESULT 9
 ADB80503
 ID ADB80503 standard; cDNA; 927 BP.
 AC ADB80503;
 XX
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulneryary; gene therapy; vaccine; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO2003042661-A2.
 XX
 XX 22-MAY-2003.
 PD
 XX
 XX 13-NOV-2002; 2002WO-US036810.
 XX
 XX 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 XX (E0SB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 XX WPI; 2003-468649/44.
 DR P-PSDB; ADB80504.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO 41; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a

CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 16.5 Length: 927
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 11 Gaps: 0

US-10-774-176-23 (1-9) x ADN38723 (1-927)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 DB 709 TCCCTGCAACCTTTATGTCTTCCTG 735

RESULT 10

AAD56198
 ID AAD56198 standard; DNA; 973 BP.
 XX
 AC AAD56198;
 XX
 DT 07-AUG-2003 (first entry)
 XX
 DE Human LRRCAPS related DNA #5.
 XX
 KW Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX
 OS Homo sapiens.

XX
 XX WO2003035831-A2.
 XX
 PD 01-MAY-2003.
 XX
 PF 21-OCT-2002; 2002WO-US033540.
 XX
 PR 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX
 PA (EXEL-) EXELIXIS INC.
 XX
 PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX
 DR WPI; 2003-421410/39.
 XX
 PT Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 PS Example 5; Page 74-75; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a

CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA
 XX
 SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 17.4 Length: 973
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-23 (1-9) x AAD56198 (1-973)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 DB 724 TCCCTGCAACCTTTATGTCTTCCTG 750

RESULT 11

ABV99349
 ID ABV99349 standard; DNA; 1156 BP.
 XX

AC ABV99349;

DT 27-JAN-2003 (first entry)

DE Human NOV8a coding sequence.

XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
 KW neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

OS Homo sapiens.

XX WO200272771-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US007288.

XX 08-MAR-2001; 2001US-0274101P.

XX 08-MAR-2001; 2001US-0274194P.

XX 08-MAR-2001; 2001US-0274281P.

XX 09-MAR-2001; 2001US-0274322P.

XX 12-MAR-2001; 2001US-0275235P.

XX 13-MAR-2001; 2001US-0275578P.

XX 13-MAR-2001; 2001US-0275579P.

XX 14-MAR-2001; 2001US-0275601P.

XX 16-MAR-2001; 2001US-0276000P.

XX 19-MAR-2001; 2001US-0276776P.

XX 19-MAR-2001; 2001US-0276994P.

XX 20-MAR-2001; 2001US-0277239P.

XX 20-MAR-2001; 2001US-0277321P.

XX 20-MAR-2001; 2001US-0277327P.

XX 21-MAR-2001; 2001US-0277338P.

XX 22-MAR-2001; 2001US-0277791P.

XX 23-MAR-2001; 2001US-0277833P.

XX 26-MAR-2001; 2001US-0278152P.

XX 27-MAR-2001; 2001US-0278894P.

XX 27-MAR-2001; 2001US-0278999P.

XX 28-MAR-2001; 2001US-0279036P.

XX 30-MAR-2001; 2001US-0279344P.

XX 30-MAR-2001; 2001US-0279995P.

XX 30-MAR-2001; 2001US-0280233P.

PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 02-APR-2001; 2001US-0280900P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 02-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 15-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 30-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 14-NOV-2001; 2001US-0333272P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337436P.
 PR 03-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.
 (CURA-) CURAGEN CORP.
 Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 Boldog FL, Li L, Zerhusen BD, Tchernav VT, Gangolli EA, Vernet CAM;
 Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 Zhong M;
 WPI; 2002-732824/79.
 P-PSDB; ABP70071.
 New NOVX polypeptides and polynucleotides, useful for preventing,
 diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 disorders, and asthma.
 Claim 16; Page 114-115; 619pp; English.
 The present invention relates to new isolated proteins (NOVX) and their
 coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 any number from 1 to 48. The NOVX proteins and coding sequences are
 useful in the manufacture of a medicament for treating a syndrome
 associated with a human disease, preferably a NOVX-associated disorder.
 The NOVX coding sequences and proteins are useful for treating,
 preventing or diagnosing diseases such as metabolic disorders, diabetes,
 obesity, infectious disease, anorexia, cancer-associated cachexia,
 cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 disease, immune disorders, hematopoietic disorders, cardiovascular
 disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 disturbances associated with obesity, metabolic syndrome X or wasting
 disorders associated with chronic diseases or various cancers. The NOVX
 coding sequences and proteins may also be used as targets for the
 identification of small molecules that modulate or inhibit e.g.

CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods
 XX
 SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	21.1	Length:	1156
Score:	43.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-23 (1-9) x ABV99349 (1-1156)

QY 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
 Db 940 TCCTGCAACCTCTTATGTTCTCTG 966

RESULT 12

ABK87175

ID ABK87175 standard; cDNA; 1260 BP.

XX ABK87175;

XX 07-OCT-2002 (first entry)

DE cDNA encoding feline oncofetal leucine-rich glycoprotein, 5T4.

XX Feline; cat; oncofetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.

XX Felis sp.

XX Key Location/Qualifiers

FT 1..1260

FT /*tag= a

FT /product= "5T4 protein"

XX WO200238612-A2.

XX 16-MAY-2002.

XX 13-NOV-2001; 2001WO-GB005004.

XX 13-NOV-2000; 2000WO-GB004317.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Myers K, Drury N, Carroll M;

XX WPI; 2002-557449/59.

XX P-PSDB; AAU98694.

PT Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.

PS Claim 4; Page 68; 68pp; English.

XX The present invention relates to the isolation of canine and feline
 CC oncofetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also

CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes feline 5T4 protein

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores: 23.3 Length: 1260
 Pred. No.: 43.00 Matches: 9
 Score: 100.0% Conservative: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 6 Gaps: 0
 DB:

US-10-774-176-23 (1-9) x ABK87175 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 Db 1048 TCCCTGCAGACTTCATTATGCTTTCTA 1074

RESULT 13

ADB97513
 ID ADB97513 standard; DNA; 1260 BP.

XX ADB97513;

XX 04-DEC-2003 (first entry)

XX Feline 5T4 antigen DNA.

XX Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.

XX Unidentified.

XX Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"

XX WO2003068816-A1.

XX 21-AUG-2003.

XX 13-FEB-2003; 2003WO-GB000670.

XX 13-FEB-2002; 2002GB-00003419.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll M, Kingsman S, Redchenko I;

XX WPI; 2003-637141/60.

XX P-PSDB; ADB97520.

XX New major histocompatibility complex class I peptide epitopes from human
 PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.

XX Disclosure; Page 67; 73pp; English.

XX The invention relates to a novel Major Histocompatibility Complex (MHC)
 CC class I peptide epitope of the 5T4 antigen. The invention further
 CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
 CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the

CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 5T4 epitope and/or its encoding nucleic acid; a method
 CC comprising detecting the presence of the 5T4 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 5T4 epitope in conjunction with an MHC class
 CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC particularly cancer. The detection method is useful for diagnosing or
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline
 CC 5T4 antigen coding DNA of the invention.

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 23.3 Length: 1260
 Score: 43.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-23 (1-9) x ADB97513 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9

Db 1048 TCCCTGCAGACTTCATTATGCTTTCTA 1074

RESULT 14

ADB97452

ID ADB97452 standard; DNA; 1260 BP.

XX ADB97452;

XX 04-DEC-2003 (first entry)

XX DNA encoding feline 5T4 protein.

XX Gene; ds; feline; Major Histocompatibility Complex class II; MHC;
 KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.

XX Unidentified.

XX Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"

XX WO2003068815-A2.

XX 21-AUG-2003.

XX 13-FEB-2003; 2003WO-GB000618.

XX 13-FEB-2002; 2002GB-00003420.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll M, Harrop R, Kingsman S;

XX WPI; 2003-663795/62.

XX P-PSDB; ADB97455.

XX New Major Histocompatibility Complex class II peptide epitope of 5T4,
 PT useful for manufacturing a medicament for diagnosing, preventing and/or
 PT treating a disease, e.g. cancer.

XX Disclosure; Page 49; 63pp; English.

XX The invention relates to a Major Histocompatibility Complex (MHC) class
CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
CC clone has a cytostatic activity, as it is useful in manufacturing a
CC medicament for preventing and/or treating a disease, particularly cancer.
CC The methods are useful for detecting T-cells capable of specifically
CC recognising a peptide epitope in conjunction with an MHC molecule, for
CC diagnosing or monitoring the progression of a cancerous disease, or for
CC detecting the presence of a peptide or nucleic acid using an agent. The
CC MHC class II peptide epitope of the invention can be used in gene therapy
CC or as part of a vaccine. This polynucleotide sequence represents the DNA
CC coding for the feline 5T4 protein.

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores: 23.3 Length: 1260
Pred. No.: 43.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 10 Gaps: 0
DB:

US-10-774-176-23 (1-9) x ADB97452 (1-1260)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
Db 1048 TCCCTGCAGACTTCTTATGCTTCTTA 1074

RESULT 15

AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX
AC AAA27058;
XX
DT 22-AUG-2000 (first entry)
XX
XX Human 5T4 tumour-associated antigen gene.
XX
XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.

OS Homo sapiens.

XX WO200029428-A2.
PN

XX 25-MAY-2000.

XX 18-NOV-1999; 99WO-GB003859.

XX 18-NOV-1998; 98GB-00025303.

PR 27-JAN-1999; 99GB-00001739.

PR 30-JUL-1999; 99GB-00017995.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll MW, Myers KA;

XX WPI; 2000-387735/33.

XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
PT response useful in vaccinating against and in treating tumors.

XX Example 2; Page 78; 79pp; English.

XX The present sequence encodes the human 5T4 tumour-associated antigen
CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
CC carcinomas but has a highly restricted expression pattern in normal adult
CC tissues. It appears to be strongly correlated to metastasis in colorectal
CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
CC induced were inoculated with a virus expression vector containing the

CC present sequence. The 5T4 antigen was shown to be effective at eliciting
CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
CC the antigen and the antigen itself can be used to elicit an immune
CC response, preferably CTL or an antibody response in a subject
XX
SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores: 23.3 Length: 1263
Pred. No.: 43.00 Matches: 9
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 3 Gaps: 0
DB:

US-10-774-176-23 (1-9) x AAA27058 (1-1263)

Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
 |||||
Db 1051 TCCCTGCAACCTCTTATGCTTCTCTG 1077

Search completed: May 27, 2006, 10:38:33
Job time : 379.5 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds
(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-23

Perfect score: 43

Sequence: 1 SLQTSYVFL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US1074176/runat.26052006.091443.24987/app.query.fasta_1
-DB=GenEmbl -OPMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORES=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss04
-USER=US1074176 @CGN 1.1 7524 @runat.26052006.091443.24987 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAFN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb env.*
- 2: gb_pat.*
- 3: gb_ph.*
- 4: gb_pl.*
- 5: gb_pr.*
- 6: gb_ro.*
- 7: gb_sts.*
- 8: gb_sy.*
- 9: gb_un.*
- 10: gb_vl.*
- 11: gb_ov.*
- 12: gb_htg.*
- 13: gb_in.*
- 14: gb_om.*
- 15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	65	2	CQ559387 Sequence
2	43	100.0	290	2	CQ687716 Sequence
3	43	100.0	475	2	CQ920916 Sequence

4	43	100.0	901	2	BD249733
5	43	100.0	901	2	AX025013
6	43	100.0	901	2	AX316088
7	43	100.0	927	2	AX829164
8	43	100.0	1156	2	DD161112
9	43	100.0	1260	2	AX467373
10	43	100.0	1260	2	AX821533
11	43	100.0	1260	2	AX821548
12	43	100.0	1263	2	BD249731
13	43	100.0	1263	2	AX025011
14	43	100.0	1263	2	AX149553
15	43	100.0	1263	2	AX316086
16	43	100.0	1263	2	AX467371
17	43	100.0	1281	2	BD249732
18	43	100.0	1281	2	AX025012
19	43	100.0	1281	2	AX316087
20	43	100.0	2053	2	DD174290
21	43	100.0	2053	5	HS5T40A
22	43	100.0	2201	11	CR855786
23	43	100.0	2333	6	AF063939
24	43	100.0	2359	2	BD127282
25	43	100.0	2359	2	CQ782724
26	43	100.0	2359	5	AK074786
27	43	100.0	2361	2	AX961916
28	43	100.0	2361	2	BD127283
29	43	100.0	2361	2	CQ782726
30	43	100.0	2361	5	AK074790
31	43	100.0	2361	6	BC087011
32	43	100.0	2379	5	BC037161
33	43	100.0	2423	6	BC058198
34	43	100.0	2527	2	AX961912
35	43	100.0	2557	2	AX961914
36	43	100.0	2714	5	AB168308
37	43	100.0	5551	5	HSA012159
38	43	100.0	7942	6	MMU012160
39	43	100.0	121909	5	HSJ492P14
C 40	43	100.0	167046	6	AC158516
C 41	43	100.0	210237	12	AC128294
C 42	43	100.0	239076	12	AC106962
C 43	43	100.0	243801	12	AC172361
44	40	93.0	1713	11	BC099984
45	40	93.0	1765	11	AB097825

ALIGNMENTS

RESULT 1	CQ559387	Sequence 29022 from Patent WO0210449.	65 bp	DNA	linear	PAT 30-JAN-2004
LOCUS	CQ559387					
DEFINITION	CQ559387					
ACCESSION	CQ559387					
VERSION	CQ559387.1	GI:41525814				
KEYWORDS						
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;					
	Sciurognathi; Muroidea; Muridae; Murinae; Mus.					
REFERENCE	1	Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.				
AUTHORS		Oligonucleotide library for detecting rna transcripts and splice				
TITLE		variants that populate a transcriptome				
JOURNAL		Patent: WO 0210449-A 29022 07-FEB-2002;				
		Compugen Inc. (US)				
FEATURES	source	Location/Qualifiers				
		1..65				
		/organism="Mus musculus"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:10090"				

ORIGIN

Alignment Scores: 2.6 Length: 65
Pred. No.: 65


```

DB:                2          Gaps:          0
US-10-774-176-23 (1-9) x BD249733 (1-901)
Qy  1 SerLeuGlnThrSerTyrValPheLeu 9
Db  608 TCCCTGCAGACTTCTTATGCTTCCTA 634

RESULT 5
AX025013
LOCUS          AX025013          901 bp      DNA          linear      PAT 15-SEP-2000
DEFINITION     Sequence 3 from Patent WO0029428.
ACCESSION      AX025013
VERSION        AX025013.1 GI:10184934
KEYWORDS
SOURCE         Canis sp.
ORGANISM       Canis sp.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
               Canis.
REFERENCE      1
AUTHORS        Carroll,M.W. and Myers,K.A.
TITLE          Polypeptide
JOURNAL        CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
               BIOMEDICA LTD (GB)
FEATURES       source
               1..901
               /organism="Canis sp."
               /mol_type="unassigned DNA"
               /db_xref="taxon:9616"
ORIGIN
Alignment Scores:
Pred. No.:      25.4      Length:      901
Score:          43.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             2          Gaps:      0

US-10-774-176-23 (1-9) x AX025013 (1-901)
Qy  1 SerLeuGlnThrSerTyrValPheLeu 9
Db  608 TCCCTGCAGACTTCTTATGCTTCCTA 634

RESULT 6
AX316088
LOCUS          AX316088          901 bp      DNA          linear      PAT 14-DEC-2001
DEFINITION     Sequence 3 from Patent EP1160323.
ACCESSION      AX316088
VERSION        AX316088.1 GI:17899280
KEYWORDS
SOURCE         Canis sp.
ORGANISM       Canis sp.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
               Canis.
REFERENCE      1
AUTHORS        Carroll,M.W. and Myers,K.A.
TITLE          5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL        Patent: EP 1160323-A 3 05-DEC-2001;
               Oxford Biomedica (UK) Limited (GB)
FEATURES       source
               1..901
               /organism="Canis sp."
               /mol_type="unassigned DNA"
               /db_xref="taxon:9616"
ORIGIN
Alignment Scores:
Pred. No.:      25.4      Length:      901
Score:          43.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             2          Gaps:      0

US-10-774-176-23 (1-9) x AX025013 (1-901)
Qy  1 SerLeuGlnThrSerTyrValPheLeu 9
Db  608 TCCCTGCAGACTTCTTATGCTTCCTA 634

RESULT 7
AX829164
LOCUS          AX829164          927 bp      DNA          linear      PAT 12-DEC-2003
DEFINITION     Sequence 57 from Patent WO02059377.
ACCESSION      AX829164
VERSION        AX829164.1 GI:39838931
KEYWORDS
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homnidae; Homo.
REFERENCE      1
AUTHORS        Mack,D.H., Gish,K.C. and Afar,D.
TITLE          Methods of diagnosis of breast cancer, compositions and methods of
               screening for modulators of breast cancer
JOURNAL        Patent: WO 02059377-A 57 01-AUG-2002;
               EOS Biotechnology, Inc. (US)
FEATURES       source
               1..927
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      26       Length:      927
Score:          43.00     Matches:      9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match:    100.0% Indels:      0
DB:             2          Gaps:      0

US-10-774-176-23 (1-9) x AX316088 (1-901)
Qy  1 SerLeuGlnThrSerTyrValPheLeu 9
Db  608 TCCCTGCAGACTTCTTATGCTTCCTA 634

RESULT 8
DD161112
LOCUS          DD161112          1156 bp     DNA          linear      PAT 23-NOV-2005
DEFINITION     Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
               Encoding The Antigens, and Methods of Use.
ACCESSION      DD161112
VERSION        DD161112.1 GI:83967439
KEYWORDS
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homnidae; Homo.
REFERENCE      1 (bases 1 to 1156)
AUTHORS        Padigaru,M., Shenoy,S.G., Pochart,P.F., Kekuda,R., Gusev,V.Y.,
               Zhong,M., Jr,R.J.T., Casman,S.J., Li,L., Miller,C.E.,
               Patturajan,M., Anderson,D.W., Malyankar,U.M., Voss,E.Z.,
               Spaderna,S.K., Gorman,L., Spytek,K.A., Liu,X., Burgess,C.E.,
               Pena,C.F.A., Gerlach,V., Smithson,G., Mezes,P.D., Rastelli,L.,
               Boldog,F.L., Guo,X., Vernet,C.A.M., Gangolli,E.A., Tchernev,V.T.
               and Zerhusen,B.D.
TITLE          Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids

```

```

JOURNAL
Encoding The Antisense, and Methods of Use
Patent: JP 200508604-A 23 07-APR-2005;
Muralidhara Padigaru,Suresh Shenoy,Remesh Kekuda,Vladimir Gusev,
Pascale Pochart,Mei Zhong,Luca Rastelli,Peter Mezes,Glenn
Smithson,XiaoJia Guo,Valerie Gerlach,Stacie Casman,Remenc
Boldog,Li Li,Bryan Zetshuen,Velizar Tcherven,Esha Gangolli,Corine
Varnet,Carol Pena,Catherine Burgess,Xiaohong Liu,Kimberly
Spytek,Linda Gorman,Steven Spaderna,Edward Voss,Uriel
Malyankar,David Anderson,Meera Patturajan,Charles Miller,Raymond J
Taupier Jr
OS Homo sapiens
PN JP 200508604-A/23
PD 07-APR-2005
PR 19-JUN-2001 US 60/299310,18-JUN-2001 US 60/299027, PR
PR 31-MAY-2001 US 60/294889,31-MAY-2001 US 60/294899, PR
30-MAY-2001 US 60/294885,09-MAR-2001 US 60/274849, PR
13-MAR-2001 US 60/275579,03-MAY-2001 US 60/288342, PR
02-MAY-2001 US 60/288066,30-APR-2001 US 60/287424, PR
13-APR-2001 US 60/283675,04-APR-2001 US 60/281194, PR
02-APR-2001 US 60/280802,02-APR-2001 US 60/280822, PR
03-MAY-2001 US 60/288528,15-MAY-2001 US 60/291190, PR
16-MAY-2001 US 60/291240,16-MAY-2001 US 60/291099, PR
12-MAR-2001 US 60/275235,08-MAR-2001 US 60/274101, PR
08-MAR-2001 US 60/274281,08-MAR-2001 US 60/274322, PR
08-MAR-2001 US 60/274194,02-APR-2001 US 60/280900, PR
30-MAR-2001 US 60/280233,30-MAR-2001 US 60/279995, PR
20-MAR-2001 US 60/277327,20-MAR-2001 US 60/277338, PR
19-MAR-2001 US 60/276994,16-MAR-2001 US 60/276776, PR
14-MAR-2001 US 60/276000,13-MAR-2001 US 60/275601, PR
13-MAR-2001 US 60/275578,20-MAR-2001 US 60/277239, PR
20-MAR-2001 US 60/277321,21-MAR-2001 US 60/277791, PR
22-MAR-2001 US 60/277833,23-MAR-2001 US 60/278152, PR
26-MAR-2001 US 60/278894,27-MAR-2001 US 60/278999, PR
27-MAR-2001 US 60/279036,28-MAR-2001 US 60/279344, PR
19-JUN-2001 US 60/299303,10-JUL-2001 US 60/304354, PR
31-JUL-2001 US 60/309198,03-DEC-2001 US 60/337426, PR
03-DEC-2001 US 60/338092,21-NOV-2001 US 60/332094, PR
14-NOV-2001 US 60/332272,14-NOV-2001 US 60/332271, PR
14-NOV-2001 US 60/332272,14-NOV-2001 US 60/332172, PR
18-NOV-2001 US 60/333184,31-OCT-2001 US 60/335301, PR
18-OCT-2001 US 60/330380,27-SEP-2001 US 60/325430, PR
27-SEP-2001 US 60/325681,12-SEP-2001 US 60/321870, PR
10-SEP-2001 US 60/318462,03-JAN-2002 US 60/345705, PR
04-DEC-2001 US 60/337185,08-MAR-2002 US 10/093463, PR
16-AUG-2001 US 60/312903
PI muralidhara padigaru,suresh g shenoy,pascale f-g pochart, PI
remesh kekuda,
PI vladimir y gusev,mei zhong,raymond j taupier jr,stacie j PI
casman,li li,
PI charles e miller,meera patturajan,david w anderson,uriel m PI
PI edward z voss,steven k spaderna,linda gorman,kimberly PI a
PI xiaohong liu,catherine e burgess,carol e a pena,valerie PI
gerlach,
PI glenn smithson,peter d mezes,luca rastelli,ferenc l boldog,
PI xiaoJia guo,
PI corine a m vernet,esha a gangolli,velizar t tchernev,bryan d
PI zetshuen
CC
FH Key Location/Qualifiers
FT CDS (1)..(1150).
FEATURES
source
1..1156
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 31.5 Length: 1156
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 2 Indels: 0
DB: 2 Gaps: 0
US-10-774-176-23 (1-9) x DD161112 (1-1156)
QY 1 SerLeuGlnThrSerTyrValPheLeu 9
DB 940 TCCTGCAACCTCTTATGTCCTCTG 966
RESULT 9
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
Felis sp.
Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"
ORIGIN
Alignment Scores:
Pred. No.: 33.9 Length: 1260
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 2 Indels: 0
DB: 2 Gaps: 0
US-10-774-176-23 (1-9) x AX467373 (1-1260)
QY 1 SerLeuGlnThrSerTyrValPheLeu 9
DB 1048 TCCTGCAGACTTCTTATGTCCTTCTA 1074
RESULT 10
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1
AUTHORS Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
TITLE MHC class I peptide epitopes from the human St4 tumor-associated
antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN

```

```
Alignment Scores:
Pred. No.: 33.9 Length: 1260
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-23 (1-9) x AX821533 (1-1260)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1048 TCCCTGCAGACTTCTTATGTCCTTCTA 1074

RESULT 11
LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE Felis catus (cat)
ORGANISM Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE 1
AUTHORS Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
TITLE MHC class II peptide epitope of 584 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
Location/Qualifiers
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"

ORIGIN
Alignment Scores:
Pred. No.: 33.9 Length: 1260
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-23 (1-9) x AX821548 (1-1260)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1048 TCCCTGCAGACTTCTTATGTCCTTCTA 1074

RESULT 12
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1263)
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002

PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K14/065,
C07K19/00,
PC C12N15/00,
CC Polypeptide
FH Key Location/Qualifiers
FT source
1..1263
/organism="Homo sapiens (human)"
/ncbi_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 34 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-23 (1-9) x BD249731 (1-1263)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCACACCTTCTTATGTCCTTCTG 1077

RESULT 13
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO0029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source
Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 34 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-23 (1-9) x AX025011 (1-1263)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCACACCTTCTTATGTCCTTCTG 1077

RESULT 14
LOCUS AX149553 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO0029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source
Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 34 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-23 (1-9) x AX025011 (1-1263)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCACACCTTCTTATGTCCTTCTG 1077

RESULT 14
AX149553
```

LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
Ellard,F.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="5T4"
ORIGIN
Alignment Scores:
Pred. No.: 34 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-23 (1-9) x AX149553 (1-1263)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCAGACTTCTTATGCTTCCTA 1077
RESULT 15
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE S24 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 34 Length: 1263
Score: 43.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-23 (1-9) x AX316086 (1-1263)
Qy 1 SerLeuGlnThrSerTyrValPheLeu 9
Db 1051 TCCCTGCAGACTTCTTATGCTTCCTG 1077

Search completed: May 27, 2006, 19:35:23
Job time : 3359.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-22

Perfect score: 46

Sequence: 1 FLFLPRDLL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Dgapop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10774176/runat_26052006_091441_24976/app_query.fasta.1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40 cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10774176 @CGN 1.1 2389 @runat_26052006_091441_24976 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	46	100.0	1281	3	AAA27059	Mouse 574
2	46	100.0	2557	12	ADI26160	Adi26160 Human cDN
3	46	100.0	2557	12	ADI26158	Adi26158 Human cDN

4	43	93.5	1260	6	ABK87175	Abk87175 cDNA enco
5	43	93.5	1260	10	ADB97513	Adb97513 Feline 5T
6	43	93.5	1260	10	ADB97452	Adb97452 DNA encod
7	41	89.1	655	13	ADG56661	Adg56661 Novel can
c 8	40	87.0	544	6	ABN60734	Abn60734 Human can
c 9	40	87.0	563	14	ACL58953	Actl58953 Human col
10	40	87.0	927	6	ABT07721	Abt07721 Breast ca
11	40	87.0	927	8	ABX76333	Abx76333 Lung canc
12	40	87.0	927	10	ADB80503	Adb80503 Ovarian c
13	40	87.0	927	11	ADN38723	Adn38723 Cancer/an
14	40	87.0	973	8	AAD56198	Aad56198 Human LRR
15	40	87.0	1156	6	ABV99349	Abv99349 Human NOV
16	40	87.0	1263	3	AAA27058	Aaa27058 Human 574
17	40	87.0	1263	4	AAR99736	Aar99736 Nucleotid
18	40	87.0	1263	6	ABK87174	Abk87174 cDNA enco
19	40	87.0	1331	8	AAD56199	Aad56199 Human LRR
c 20	40	87.0	1441	3	AAAC39040	Aac39040 Arabidops
c 21	40	87.0	1489	13	ADT17460	Adt17460 Plant CDN
c 22	40	87.0	2020	10	ADJ56299	Adj56299 Human CDN
23	40	87.0	2053	8	ACC51052	Acc51052 Human bla
24	40	87.0	2053	8	ABX76332	Abx76332 Lung canc
25	40	87.0	2053	8	AAD56197	Aad56197 Human LRR
26	40	87.0	2053	8	AAD56200	Aad56200 Human LRR
27	40	87.0	2053	11	ADN38721	Adn38721 Cancer/an
28	40	87.0	2053	12	ADL06473	Adl06473 Human tum
29	40	87.0	2053	12	ADN03961	Adn03961 Antipsori
30	40	87.0	2053	13	ADR25444	Adr25444 Breast ca
31	40	87.0	2053	13	ACN38510	Acn38510 Tumour-as
32	40	87.0	2053	13	ADV35098	Adv35098 Human cDN
33	40	87.0	2053	14	AED17761	Aed17761 Fibrotic
34	40	87.0	2338	5	AAS87175	Aas87175 DNA encod
35	40	87.0	2359	4	AAK94253	Aak94253 Human ful
36	40	87.0	2359	12	ADL30831	Adl30831 Full leng
37	40	87.0	2361	4	AAK94254	Aak94254 Human ful
38	40	87.0	2361	12	ADI26162	Adi26162 Human cDN
39	40	87.0	2361	12	ADL30833	Adl30833 Full leng
c 40	39	84.8	371	4	AAL34954	Aal34954 Human mus
c 41	39	84.8	371	8	ABX57942	Abx57942 cDNA enco
c 42	39	84.8	371	12	ADJ27669	Adj27669 Human mus
c 43	39	84.8	425	2	AAK41515	Aak41515 Human sec
c 44	39	84.8	624	13	ADX48968	Adx48968 Plant ful
c 45	39	84.8	626	3	AAK75574	Aak75574 Human ORF

ALIGNMENTS

RESULT 1	
AAA27059	AAA27059 standard; DNA; 1281 BP.
AC	AAA27059;
XX	
DT	22-AUG-2000 (first entry)
XX	
DE	Mouse 574 tumour-associated antigen gene.
XX	
KW	Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW	immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW	ds.
OS	Mus musculus.
XX	
PN	WO200029428-A2.
XX	
PD	25-MAY-2000.
XX	
PF	18-NOV-1999; 99WO-GE003859.
XX	
PR	18-NOV-1998; 98GB-00025303.
PR	27-JAN-1999; 99GB-00001739.
PR	30-JUL-1999; 99GB-00017995.
XX	
PA	(OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll MW, Myers KA;
 XX WPI; 2000-387735/33.
 XX
 PT Tumor associated antigen, ST4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.
 XX
 PS Example 2; Page 78; 79pp; English.
 XX
 CC The present sequence encodes the mouse ST4 tumour-associated antigen
 CC (TAA). The TAA ST4 is a glycoprotein which is widely expressed in
 CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. ST4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The ST4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject. The
 CC present sequence appears in GenBank at accession number AJ012160
 XX
 SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 31 Length: 1281
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-22 (1-9) x AAA27059 (1-1281)
 QY 1 PheLeuPheLeuProArgAspLeuLeu 9
 Db 682 TTTCTTTTCTGCTCGGGACTTACTA 708
 RESULT 2
 ADI26160
 ID ADI26160 standard; cDNA; 2557 BP.
 AC ADI26160;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes SPAT6 activation #63.
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 PF 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00337505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 PA (ASAH) ASahi KASEI KK.
 XX

PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX WPI; 2004-122214/12.
 DR P-PSDB; ADI26161.
 XX
 PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 PS Claim 4; SEQ ID NO 125; 1368pp; English.
 XX
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis, AIDS. The protein has effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 67.9 Length: 2557
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-22 (1-9) x ADI26160 (1-2557)
 QY 1 PheLeuPheLeuProArgAspLeuLeu 9
 Db 1237 TTTCTTTTCTGCTCGGGACTTACTA 1263
 RESULT 3
 ADI26158
 ID ADI26158 standard; cDNA; 2557 BP.
 XX
 AC ADI26158;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #62.
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.

XX PF 05-JUN-2003; 2003WO-JP007123.
 XX PR 05-JUN-2002; 2002JP-00164257.
 XX PR 06-JUN-2002; 2002US-0385912P.
 XX PR 26-DEC-2002; 2002JP-00377326.
 XX PR 27-DEC-2002; 2002US-0436467P.
 XX PR 15-MAY-2003; 2003JP-00137505.
 XX PR 16-MAY-2003; 2003US-0470836P.
 XX FA (ASAH) ASAH KASEI KK.
 XX XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX WPI; 2004-122214/12.
 XX DR P-PSDB; ADI26159.
 XX XX New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX XX Claim 4; SEQ ID NO 123; 1368pp; English.
 XX XX The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX XX Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;
 SQ Alignment Scores: 67.9 Length: 2557
 Pred. No.: 46.00 Matches: 9
 Score: 100.0% Conservative: 0
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 100.0% Indels: 0
 Query Match: 12 Gaps: 0
 DB: 1237 TTCTCTTCTGCTCGGACTTACTA 1263
 US-10-774-176-22 (1-9) x ADI26158 (1-2557)
 Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
 DB 1237 TTCTCTTCTGCTCGGACTTACTA 1263
 RESULT 4
 ABK87175
 ID ABK87175 standard; cDNA; 1260 BP.
 XX AC ABK87175;
 XX DT 07-OCT-2002 (first entry)
 XX DE cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
 XX XX

KW Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 XX foetal abnormality; foetal sex determination; gene; ss.
 OS Felis sp.
 FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "5T4 protein"
 XX WO200238612-A2.
 XX PN 16-MAY-2002.
 XX PD 13-NOV-2001; 2001WO-GB005004.
 XX PF 13-NOV-2000; 2000WO-GB004317.
 XX PR (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PA Myers K, Drury N, Carroll M;
 XX PI WPI; 2002-557449/59.
 XX DR P-PSDB; AAU98694.
 XX XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.
 XX XX Claim 4; Page 68; 68pp; English.
 PS The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes feline 5T4 protein
 XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores: 119 Length: 1260
 Pred. No.: 43.00 Matches: 8
 Score: 100.0% Conservative: 1
 Percent Similarity: 100.0% Mismatches: 0
 Best Local Similarity: 88.9% Indels: 0
 Query Match: 93.5% Gaps: 0
 DB: 6
 US-10-774-176-22 (1-9) x ABK87175 (1-1260)
 Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
 DB 661 TTCTCTTCTGCTCGGACTTACTG 687
 RESULT 5
 ADB97513
 ID ADB97513 standard; DNA; 1260 BP.
 XX AC ADB97513;
 XX XX

DT	04-DEC-2003	(first entry)
XX	Feline 5T4 antigen DNA.	
XX	Major Histocompatibility Complex class I peptide epitope; MHC;	
KW	5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;	
KW	cytostatic; cancer; feline; gene; ds.	
XX	Unidentified.	
XX		
XX	Location/Qualifiers	
PH	Key	
DE	CDS	
FT	1..1260	
FT	/+tag= a	
FT	/product= "Feline 5T4 antigen protein"	
XX		
XX	WC2003068816-A1.	
PN		
PD	21-AUG-2003.	
XX		
PF	13-FEB-2003; 2003WO-GB000670.	
PR	13-FEB-2002; 2002GB-00003419.	
PA	(OXFO-) OXFORD BIOMEDICA UK LTD.	
XX		
PI	Carroll M, Kingsman S, Redchenko I;	
XX		
DR	WPI; 2003-637141/60.	
DR	P-PSDB; ADB97520.	
XX		
PT	New major histocompatibility complex class I peptide epitopes;	
PT	5T4 tumor-associated antigen, useful for preventing and/or treating	
PT	disease, particularly cancer.	
XX		
PS	Disclosure; Page 67; 73pp; English.	
XX		
CC	The invention relates to a novel Major Histocompatibility Complex	
CC	class I peptide epitope of the 5T4 antigen. The invention further	
CC	provides a polypeptide string comprising the 5T4 epitope; a nucleic	
CC	sequence encoding the 5T4 epitope or a polypeptide string of the	
CC	epitope; a vector system capable of delivering the 5T4 epitope	
CC	acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide	
CC	5T4 epitope, its encoding nucleic acid, or the vector system; and	
CC	comprising the above; a method for treating and/or preventing	
CC	in a subject by administering the vaccine; an agent capable of	
CC	specifically to the 5T4 epitope and its encoding nucleic acid;	
CC	comprising detecting the presence of the 5T4 epitope or its	
CC	nucleic acid in a subject; and a T cell line or clone capable of	
CC	specifically recognising the 5T4 epitope in conjunction with an	
CC	I molecule. The 5T4 epitope has cytostatic activity. The vaccine	
CC	comprising the 5T4 epitope or its encoding nucleic acid and the	
CC	system or cell is useful in the prevention and/or treatment of	
CC	particularly cancer. The detection method is useful for diagnosis	
CC	monitoring the progression of a cancerous disease, and for determining	
CC	presence of the 5T4 epitope or its nucleic acid. The T cell line	
CC	is useful in the manufacture of a medicament for treating and/or	
CC	preventing a disease. This polynucleotide sequence represents	
CC	5T4 antigen coding DNA of the invention.	
XX		
SQ	Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;	
Alignment Scores:		
Pred. No.:	119 Length: 1260	
Score:	43.00 Matches: 8	
Percent Similarity:	100.0% Conservative: 1	
Best Local Similarity:	88.9% Mismatches: 0	
Query Match:	93.5% Indels: 0	
DB:	10 Gaps: 0	
US-10-774-176-22 (1-9) x ADB97513 (1-1260)		
Qy	1 PheLeuPheLeuProArgAspLeuLeu 9 ::	

Db 661 TTCCTCTTCTTGCCCTCGGACGTACTG 687

RESULT 6
 ADB97452
 ID ADB97452 standard; DNA; 1260 BP.
 XX
 AC ADB97452;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE DNA encoding feline 5T4 protein.
 XX
 KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
 KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
 XX
 OS Unidentified.
 XX
 EH Key Location/Qualifiers
 FT CDS 1..1260
 FT FT /*tag= a
 FT FT /product= "Feline 5T4 antigen protein".
 XX
 PN WO2003068815-A2.
 XX
 PD 21-AUG-2003.
 XX
 PF 13-FEB-2003; 2003WO-GB000618.
 XX
 PR 13-FEB-2002; 2002GB-00003420.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Carroll M, Harrop R, Kingsman S;
 XX
 DR WPI; 2003-663795/62.
 DR P-PSDB; ADB97455.
 XX
 PT New Major Histocompatibility Complex class II peptide epitope of 5T4,
 PT useful for manufacturing a medicament for diagnosing, preventing and/or
 PT treating a disease, e.g. cancer.
 XX
 PS Disclosure; Page 49; 63pp; English.
 XX
 CC The invention relates to a Major Histocompatibility Complex (MHC) class
 CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
 CC clone has a cytostatic activity, as it is useful in manufacturing a
 CC medicament for preventing and/or treating a disease, particularly cancer
 CC The methods are useful for detecting T-cells capable of specifically
 CC recognising a peptide epitope in conjunction with an MHC molecule, for
 CC diagnosing or monitoring the progression of a cancerous disease, or for
 CC detecting the presence of a peptide or nucleic acid using an agent. The
 CC MHC class II peptide epitope of the invention can be used in gene therapy
 CC or as part of a vaccine. This polynucleotide sequence represents the DNA
 CC coding for the feline 5T4 protein.
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 119 Length: 1260
 Score: 43.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 93.5% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-22 (1-9) x ADB97452 (1-1260)

Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
 |||||
 Db 661 TTCCTCTTCTTGCCCTCGGACGTACTG 687

RESULT 7
 ADQ56661

ID ADQ56661 standard; DNA; 655 BP.
AC ADQ56661;
XX
DT 21-OCT-2004 (first entry)
XX
DE Novel canine microarray-related DNA sequence SeqID7963.
XX
KW canine microarray; drug screening; toxicity assay;
KW environmental pollutant; cellular response; gene expression profile;
KW toxic response; liver necrosis; fatty liver disease;
KW protein adduct formation; hepatitis; dog; ds.
XX
OS Canis familiaris.
XX
FN WO2004063324-A2.
XX
XX 29-JUL-2004.
XX
XX 05-MAY-2003; 2003WO-US013853.
XX
XX 03-MAY-2002; 2002US-0377240P.
XX
XX (GENE-) GENE LOGIC INC.
XX (PFIZ) PFIZER PROD INC.
XX
XX Diggans JC, Porter M, Wei T;
XX
XX WPI; 2004-561890/54.
XX
XX New isolated nucleic acid molecule, useful for drug screening and
XX toxicity assays or for assessing the impact, including toxicity, of a
XX compound, pharmaceutical agent or environmental pollutant on a cell or
XX living organism.
XX
XX Claim 1; SEQ ID NO 7963; 41pp; English.
XX
XX This invention is related to a novel isolated canine nucleic acid
XX sequences and the construction of canine microarrays containing a
XX significant portion of the canine genome. The isolated canine nucleic
XX acid sequences of the invention may be useful for drug screening and
XX toxicity assays. The invention is therefore useful for assessing the
XX impact, including toxicity, of a compound, pharmaceutical agent or
XX environmental pollutant on a cell or living organism. The methods are
XX useful for detecting genes that are up- or down-regulated in canines in a
XX disease state. The sequences are useful as diagnostic agents or markers
XX to detect a cellular response in a sample individually or as part of a
XX gene expression profile. It is also useful as a target for agents that
XX modulate gene expression or activity. The database is useful for
XX producing electronic Northern blots that allow the user to determine the cell
XX type or tissue in which a given gene is expressed and to allow
XX determination of the abundance or expression level of a given gene in a
XX particular tissue or cell. The methods are useful for determining the
XX similarity of a toxic response to one or more individual compounds. The
XX methods are useful for predicting at least one toxic response or the
XX likelihood that a compound or test agent will induce various specific
XX pathologies such as those of the liver (liver necrosis, fatty liver
XX disease, protein adduct formation or hepatitis), those of the kidney,
XX heart, brain or testes, or other pathologies associated with at least one
XX of the toxins. The methods are also useful for predicting or elucidating
XX the potential cellular pathways influenced, induced or modulated by the
XX compound or test agent due to the similarity of the expression profile
XX compared to the profile induced by a known toxin. The present sequence is
XX that of a canine DNA sequence which was claimed for use during the
XX production of a canine microarray of the invention.

SEQ Sequence 655 BP; 174 A; 89 C; 131 G; 253 T; 0 U; 8 Other;
Alignment Scores:
Pred. No.: 141 Length: 655
Score: 41.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1

Query Match: 89.1% Indels: 0
DB: 13 Gaps: 0
US-10-774-176-22 (1-9) x ADQ56661 (1-655)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
DB 589 TTCTTGTCTTACCACGTGATTGTTA 615
RESULT 8
ABN60734/c
ID ABN60734 standard; cDNA; 544 BP.
XX
XX AC ABN60734;
XX
XX DT 28-JUN-2002 (first entry)
XX
XX DE Human cancer related polynucleotide SEQ ID NO 701.
XX
XX KW Human; cytostatic; gene expression; gene mapping; tissue profiling;
XX KW gene therapy; cancer; tumour; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200214500-A2.
XX
XX PD 21-FEB-2002.
XX
XX PF 16-AUG-2001; 2001WO-US025840.
XX
XX PR 16-AUG-2000; 2000US-0226326P.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
XX Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX WPI; 2002-241905/29.
XX
XX New nucleic acid for producing a polypeptide, detecting differentially
XX expressed genes correlated with a cancerous state of a mammalian cell,
XX PT and inhibiting tumor growth.
XX
XX Claim 1; SEQ ID NO 701; 883pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
XX with cytostatic activity. The polynucleotide is used to produce a
XX polypeptide, to detect differentially expressed genes correlated with a
XX cancerous state of a mammalian cell and to inhibit tumour growth. The
XX polynucleotide is used as a probe in mapping and tissue profiling. The
XX encoded polypeptide and antibodies to the polypeptide can also be used
XX for therapeutic and diagnostic purposes. The polynucleotide is useful for
XX gene therapy. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 544 BP; 189 A; 115 C; 126 G; 114 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 180 Length: 544
Score: 40.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 87.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-22 (1-9) x ABN60734 (1-544)
QY 2 LeuPheLeuProArgAspLeuLeu 9
DB 76 CTGTTTTTGCACGACCTGCTT 53

RESULT 9
 ACL58953/c
 ID ACL58953 standard; cDNA; 563 BP.
 AC
 AC ACL58953;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE Human colon cancer differentially expressed polynucleotide, SEQ ID:5088.
 XX
 DE Differential expression; diagnosis; therapy; drug screening; cancer;
 KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200500087-A2.
 XX
 PD 06-JAN-2005.
 XX
 XX 13-MAY-2004; 2004WO-US015421.
 XX
 PR 03-JUN-2003; 2003US-0475872P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Randazzo F, Moler E, Escobedo J, Garcia PD;
 XX
 DR WPI; 2005-075421/08.
 XX
 XX New isolated polynucleotides, which are differentially expressed in colon
 PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
 PT cancer, or pancreatic cancer.
 XX
 PS Claim 1; SEQ ID NO 5088; 97pp; English.
 XX
 CC The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
 CC are differentially expressed in colon cancer cells. The invention also
 CC relates to vectors and host cells comprising a differentially expressed
 CC polynucleotide of the invention; a method for detecting a cancerous cell
 CC by detection of a gene product of the polynucleotides; a method for
 CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
 CC of the polynucleotides; a method of treating an individual with cancer by
 CC administration of a modulator of a gene product of the polynucleotides;
 CC and an isolated antibody that specifically binds to a polypeptide encoded
 CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
 CC antibodies, and methods are useful for the detection of cancerous cells;
 CC for the diagnosis, prognosis and management of cancer; for the
 CC identification of agents that modulate the phenotype of cancerous cells;
 CC for the identification of therapeutic targets for cancer chemotherapy;
 CC and for the treatment of cancer, especially colon cancer and metastasized
 CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
 CC are also useful as a source of probes or primers for use in diagnostic
 CC methods. The differentially expressed polynucleotides or their encoded
 CC proteins can additionally be used as vaccines to modulate primary immune
 CC responses for the prevention or treatment of cancer. The present sequence
 CC represents a specifically claimed polynucleotide which is differentially
 CC expressed in colon cancer. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 563 BP; 192 A; 119 C; 135 G; 117 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 187 Length: 563
 Score: 40.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 14 Indels: 0
 DB: 17 Gaps: 0

US-10-774-176-22 (1-9) x ACL58953 (1-563)

QY 2 LeuPheLeuProArgAspLeuLeu 9
 DB 76 CTGTTTTCCTGCAAGACACCTGCTT 53

RESULT 10

ABT07721
 ID ABT07721 standard; DNA; 927 BP.
 AC
 AC ABT07721;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Breast cancer-associated gene sequence 29.
 XX
 DE Gene; ds; breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.
 XX
 OS Unidentified.
 XX
 PN WO200259377-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 24-JAN-2002; 2002WO-US002242.
 XX
 PR 24-JAN-2001; 2001US-0263965P.
 PR 02-FEB-2001; 2001US-0265928P.
 PR 09-APR-2001; 2001US-00829472.
 PR 09-APR-2001; 2001US-0282698P.
 PR 04-MAY-2001; 2001US-0288590P.
 PR 29-MAY-2001; 2001US-0294443P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Mack DH, Gish KC, Afar D;
 DR WPI; 2002-583738/62.
 DR N-PSDB; ABJ05564.
 XX
 PT Detecting a breast cancer-associated transcript in a patient's cell,
 PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.
 XX
 PS Claim 9; Page 372; 414pp; English.
 XX
 CC The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
 CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 329 Length: 927
 Score: 40.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 87.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-22 (1-9) x ABT07721 (1-927)		Alignment Scores:		Pred. No.: 329		Length: 927	
Qy	1 PheLeuPheLeuProArgAspLeuLeu 9	Score:	40.00	Matches:	7		
Db	322 TTCCTTTACCTGCCGCGGATGTGCTG 348	Percent Similarity:	100.0%	Conservative:	2		
		Best Local Similarity:	77.8%	Mismatches:	0		
		Query Match:	87.0%	Indels:	0		
		DB:	8	Gaps:	0		
RESULT 11		US-10-774-176-22 (1-9) x ABX76333 (1-927)					
ABX76333	ID ABX76333 standard; DNA; 927 BP.						
XX	AC ABX76333;						
XX	DT 02-APR-2003 (first entry)						
XX	DE Lung cancer-associated polynucleotide #197.						
XX	KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;						
XX	KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;						
XX	KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;						
XX	KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;						
XX	KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.						
XX	OS Unidentified.						
XX	OS DE						
XX	PN WO200286443-A2.						
XX	PD 31-OCT-2002.						
XX	PF 18-APR-2002; 2002WO-US012476.						
XX	PR 18-APR-2001; 2001US-0284770P.						
XX	PR 10-MAY-2001; 2001US-0290492P.						
XX	PR 09-NOV-2001; 2001US-0339245P.						
XX	PR 13-NOV-2001; 2001US-0350666P.						
XX	PR 29-NOV-2001; 2001US-0334370P.						
XX	PR 12-APR-2002; 2002US-0372246P.						
XX	PA (EOSB-) EOS BIOTECHNOLOGY INC.						
XX	PI Aziz N, Murray R;						
XX	DR WPI; 2003-093161/08.						
XX	DR P-PSDB; ABUS6604.						
PT	PT Detecting a lung cancer-associated transcript in a cell from a patient						
PT	PT for treating lung cancer, by contacting a biological sample from the						
PT	PT patient with a polynucleotide that exhibits increased or decreased						
PT	PT expression in lung cancer.						
XX	XX Claim 22; Page 336; 453pp; English.						
XX	XX The invention relates to a method for detecting a lung cancer-associated						
XX	XX transcript in a cell from a patient, comprising contacting a biological						
XX	XX sample from the patient with a polynucleotide that selectively hybridizes						
XX	XX to a sequence that is at least 80 % identical to a gene that exhibits						
XX	XX increased or decreased expression in lung cancer samples. Lung cancer-						
XX	XX associated polynucleotides and polypeptides are used for identifying a						
XX	XX compound that modulates a lung cancer-associated polypeptide, for						
XX	XX inhibiting proliferation of a lung cancer-associated cell to treat lung						
XX	XX cancer in a patient and for treating a mammal having lung cancer by						
XX	XX administering a modulatory compound identified. The methods are useful						
XX	XX for treating lung cancer, such as small cell lung cancer, non-small cell						
XX	XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,						
XX	XX emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,						
XX	XX hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and						
XX	XX bronchiectasis. The genes, polynucleotides and polypeptides are useful						
XX	XX for diagnostic purposes and as targets for screening for therapeutic						
XX	XX compounds that modulate lung cancer, such as antibodies. Sequences						
XX	XX ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the						
XX	XX invention						
SQ	SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;						

XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 329 Length: 927
Score: 40.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 87.0% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-22 (1-9) x ADB80503 (1-927)
Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 322 TTCCTTTACCTCCGCGGATGTGCTG 348
RESULT 13
ADN38723
ID ADN38723 standard; cDNA; 927 BP.
AC ADN38723;
XX
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
XX
DE Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2003042661-A2.
XX
XX 22-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036810.
XX
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-039775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468645/44.
DR P-PSDB; ADN38724.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

XX Claim 8; SEQ ID NO 41; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADM40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 329 Length: 927
Score: 40.00 Matches: 7
Percent Similarity: 100.0% Conservative: 2
Best Local Similarity: 77.8% Mismatches: 0
Query Match: 87.0% Indels: 0
DB: 11 Gaps: 0
US-10-774-176-22 (1-9) x ADN38723 (1-927)
Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 322 TTCCTTTACCTCCGCGGATGTGCTG 348
RESULT 14
AAD56198
ID AAD56198 standard; DNA; 973 BP.
XX
XX AAD56198;
XX
XX 07-AUG-2003 (first entry)
XX
XX Human LRRCAPS related DNA #5.
XX
XX Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX WO2003035831-A2.
XX
XX 01-MAY-2003.
XX
XX 21-OCT-2002; 2002WO-US033540.
XX
XX 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357600P.
PR 01-MAR-2002; 2002US-0361196P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
XX WPI; 2003-421410/39.
XX
XX Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX

Example 5; Page 74-75; 99pp; English.

The invention relates to a method of identifying a candidate p53 pathway modulating agent. The method involves contacting an assay system comprising a purified leucine rich repeat, capricious related (LRRCAPS) polypeptide or nucleic acid or its fragment with a test agent and detecting a test agent-biased activity, where a difference between the test agent-biased activity and the reference activity identifies the test agent as a candidate p53 pathway modulating agent. The method is useful for identifying a candidate p53 pathway-modulating agent for preparing a composition for diagnosing or treating cancer. The invention is useful in gene therapy. The present sequence is human LRRCAPS related DNA

Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 348 Length: 973
 Score: 40.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 77.8% Mismatches: 0
 Query Match: 87.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-22 (1-9) x AAD56198 (1-973)

Qy 1 PheLeuPheLeuProArgAspLeuLeu 9

Db 337 TTCCTTACCTGCCGCGGATGTGCTG 363

RESULT 15

ABV9349

ID ABV9349 standard; DNA; 1156 BP.

AC ABV9349;

XX 27-JAN-2003 (first entry)

XX Human NOV8a coding sequence.

XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
 KW neutropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOV; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

XX Homo sapiens.

XX WO200272771-A2.

XX 19-SEP-2002.

XX 08-MAR-2002; 2002WO-US007288.

XX 08-MAR-2001; 2001US-0274101P.

XX 08-MAR-2001; 2001US-0274194P.

XX 08-MAR-2001; 2001US-0274281P.

XX 09-MAR-2001; 2001US-0274322P.

XX 09-MAR-2001; 2001US-0274849P.

XX 12-MAR-2001; 2001US-0275233P.

XX 13-MAR-2001; 2001US-0275578P.

XX 13-MAR-2001; 2001US-0275579P.

XX 13-MAR-2001; 2001US-0275601P.

XX 16-MAR-2001; 2001US-0276000P.

XX 19-MAR-2001; 2001US-0276776P.

XX 20-MAR-2001; 2001US-0276994P.

XX 20-MAR-2001; 2001US-0277239P.

XX 20-MAR-2001; 2001US-0277321P.

XX 20-MAR-2001; 2001US-0277327P.

PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0278999P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0279895P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 02-APR-2001; 2001US-0280900P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 02-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 03-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 30-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 21-NOV-2001; 2001US-0333722P.
 PR 03-DEC-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 04-DEC-2001; 2001US-0338092P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.
 XX (CURA-) CURAGEN CORP.

(CURA-) CURAGEN CORP.

Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 Voss EZ, Malyankar UM, Anderson DW, Patutarajan M, Miller CE;
 Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 Zhong M;

WPI; 2002-732824/79.

P-PSDB; ABP70071.

New NOVX polypeptides and polynucleotides, useful for preventing,
 diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 disorders, and asthma.

Claim 16; Page 114-115; 619pp; English.

The present invention relates to new isolated proteins (NOVX) and their
 coding sequences (ABV93327-ABV93595 and ABP70049-ABP70149), where X is
 any number from 1 to 48. The NOVX proteins and coding sequences are
 useful in the manufacture of a medicament for treating a syndrome
 associated with a human disease, preferably a NOVX-associated disorder.

CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, haematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods
 CC
 CC
 CC

SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	423	Length:	1156
Score:	40.00	Matches:	7
Percent Similarity:	100.0%	Conservative:	2
Best Local Similarity:	77.8%	Mismatches:	0
Query Match:	87.0%	Indels:	0
DB:	6	Gaps:	0

US-10-774-176-22 (1-9) x ABV99349 (1-1156)

QY	1	PheLeuPheLeuProArgAspLeuLeu 9
DB	553	TTCTTTTACCTGCCGGGAGTGCTG 579

Search completed: May 27, 2006, 10:37:43
 Job time : 382.5 secs

ECORI-NotI cut cDNA was then ligated into the vector. Vector:
pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia coli DH10B.

FEATURES

source
Location/Qualifiers
1..1060
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST360e16"
/clone_lib="CSEQCHN71"
/dev_stage="stage 36"

ORIGIN

Alignment Scores:
Pred. No.: 11.4 Length: 1060
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-22 (1-9) x BX932209 (1-1060)

Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 680 TTCCTTTTCTTCCAGAGACTGCTT 654

RESULT 2

BD249732 1281 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Polypeptide.
ACCESSION BD249732
VERSION BD249732.1 GI:33059502
KEYWORDS JP 2002530060-A/2.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL OXFORD BIOMEDICA LTD

COMMENT
OS Mus musculus (mouse)
PN JP 2002530060-A/2
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4

PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K14/065,
PC C07K19/00,
PC C12N15/00
CC Polypeptide
FH Key Location/Qualifiers
FT source 1..1281
FT source /organism="Mus musculus (mouse)"

FEATURES

source
Location/Qualifiers
1..1281
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 13 Length: 1281
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-22 (1-9) x BD249732 (1-1281)

Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 682 TTCCTTTTCTTCCCTCGGACTTACTA 708

RESULT 3

AX025012 1281 bp DNA linear PAT 15-SEP-2000
LOCUS
DEFINITION Sequence 2 from Patent WO0029428.
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL PATENT: WO 0029428-A 2 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)

FEATURES
source
Location/Qualifiers
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 13 Length: 1281
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-22 (1-9) x AX025012 (1-1281)

Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 682 TTCCTTTTCTTCCCTCGGACTTACTA 708

RESULT 4

AX316087 1281 bp DNA linear PAT 14-DEC-2001
LOCUS
DEFINITION Sequence 2 from Patent EP1160323.
ACCESSION AX316087
VERSION AX316087.1 GI:17899279
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL PATENT: EP 1160323-A 2 05-DEC-2001;
OXFORD BIOMEDICA (UK) LIMITED (GB)

FEATURES
source
Location/Qualifiers
1..1281
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 13 Length: 1281
Score: 46.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

US-10-774-176-22 (1-9) x AX316087 (1-1281)

QY 1 PheLeuPheLeuProArgAspLeuLeu 9

DB 682 TTTCTTTTCTGCTCGGACTTACTA 708

RESULT 5
 BC058198 2423 bp mRNA linear ROD 21-OCT-2003
 LOCUS Mus musculus trophoblast glycoprotein, mRNA (cdna clone MGC:68145
 DEFINITION IMAGE:5353871), complete cds.

ACCESSION BC058198
 VERSION BC058198.1 GI:34849573
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.

1 (bases 1 to 2423)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
 Mancini, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullighy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalusz, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.,
 human and mouse cDNA sequences
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2423)
 Strausberg, R.

Direct Submission
 Submitted (15-SEP-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland
 Web site: <http://www.nisac.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 123 Row: P Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

source

1..2423
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:68145 IMAGE:5353871"
 /tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..2423
 /gene="Tpbp"
 /note="synonym: 5T4"
 /db_xref="GeneID:21983"
 /db_xref="MGI:1341264"
 402..1682
 /gene="Tpbp"
 /codon_start=1
 /product="Tpbp protein"
 /protein_id="AAH58198.1"
 /db_xref="GI:34849574"
 /db_xref="GeneID:21983"
 /db_xref="MGI:1341264"
 /translation="MPGASGRSPAGDGRRLRLARLALLVLLGWVSASAPSSVPSSTNS
 PAFVAFASQAPPPAERCPACECSEARTKVCVNRNLLEVPADLPYVYRNLELTGNQ
 MTVLPAGAPQFPPLAEALNLSGNHLKEVCAFEHLPLRLDLNPLNLSAF
 APAGNASVSPSPLEELINHIIVPPEDQONGSFCGVAPEGMVAALRSGLALRGL
 TCLEASNHFLEPLDLAQLPSRLYLDLRNLSLVITVASFRNLTHLESLHLELNL
 KVLHNSLAEWGLAHVKVFLDNNPWCDYMDVMWLKETVVPDKARLTCAFPPEK
 MNRNLDLNSDDYDCAVLPQSLQTSYVFLGIVLALIGLIFLLVLYLNKRGIKKWHH
 NIRDACRDHMGHYHYRYEINADPRLTNLSNSDV"
 642..1262
 /gene="Tpbp"
 /note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
 protein [Function unknown]"
 1299..1415
 /db_xref="CDD:COG4886"
 /gene="Tpbp"
 /note="LRRCT; Region: Leucine rich repeat C-terminal
 domain"
 /db_xref="CDD:smart00082"

gene

CDS

misc_feature

misc_feature

ORIGIN

Alignment Scores:
 Pred. No.: 20.4 Length: 2423
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-22 (1-9) x BC058198 (1-2423)

QY 1 PheLeuPheLeuProArgAspLeuLeu 9

DB 1083 TTTCTTTTCTGCTCGGACTTACTA 1109

RESULT 6

AX961912

LOCUS

DEFINITION

AX961912

AX961912

AX961912.1

GI:40881322

VERSION

KEYWORDS

SOURCE

AX961912 Sequence 123 from Patent WO03104277.
 AX961912 Sequence 123 from Patent WO03104277.
 AX961912.1 GI:40881322
 Mus musculus (house mouse)

2557 bp DNA linear PAT 14-JAN-2004

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 123 18-DEC-2003;
Aashi Kasei Kabushiki Kaisha (JP)
FEATURES
source 1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881323"
/translating="MPGAGSGPSAGDGRRLRLALVLLGWVSASAPSSVPSSTTS
PAFLASGSAQPPPAECPCACESEARTVKVNRNLLLEVPADLPPTVRLNLTGNQ
MTVLPAGAFARQPPPLADEALNLSGNHLKEVCAGAFHLGLRLDLSHNPLTNSAF
VFAGSNASVAPSPLLEILNHI VPPEDORQNGFEGMVAPEGMVAALRSGALRGL
TRLEASNHFELFLPDLIAQLPSLYLDLRNNSLSVLTYSFRLTHLESLELDNAL
KVLHNSLTAEQGLAHVKVFLDNNPWCDCYMDMVAWLKETEVPDKARLTCAPEK
MNRGLDLNSSLDDCDVAVLPQSLOTYSYVFLGIVLALIGAIFLLVLYLNRRKGIKKMMH
NIRDACRDMEGHYRYEINADPRLTNLSSNDV"

ORIGIN
Alignment Scores: 21.2 Length: 2557
Pred. No.: 46.00 Matches: 9
Score: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-22 (1-9) x AX961912 (1-2557)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 1237 TTCTTTCTGCTCGGACTTACTA 1263

RESULT 7
LOCUS AX961912 2557 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 125 from Patent WO03104277.
ACCESSION AX961914
VERSION AX961914.1 GI:40881324
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 125 18-DEC-2003;
Aashi Kasei Kabushiki Kaisha (JP)
FEATURES
source 1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAF06466.1"
/db_xref="GI:40881325"
/translating="MPGAGSGPSAGDGRRLRLALVLLGWVSASAPSSVPSSTTS
PAFLASGSAQPPPAECPCACESEARTVKVNRNLLLEVPADLPPTVRLNLTGNQ
MTVLPAGAFARQPPPLADEALNLSGNHLKEVCAGAFHLGLRLDLSHNPLTNSAF
VFAGSNASVAPSPLLEILNHI VPPEDORQNGFEGMVAPEGMVAALRSGALRGL
TRLEASNHFELFLPDLIAQLPSLYLDLRNNSLSVLTYSFRLTHLESLELDNAL
KVLHNSLTAEQGLAHVKVFLDNNPWCDCYMDMVAWLKETEVPDKARLTCAPEK
MNRGLDLNSSLDDCDVAVLPQSLOTYSYVFLGIVLALIGAIFLLVLYLNRRKGIKKMMH
NIRDACRDMEGHYRYEINADPRLTNLSSNDV"

ORIGIN
Alignment Scores: 21.2 Length: 2557
Pred. No.: 46.00 Matches: 9
Score: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-22 (1-9) x AX961912 (1-2557)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 1237 TTCTTTCTGCTCGGACTTACTA 1263

RESULT 7
LOCUS AX961914 2557 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 125 from Patent WO03104277.
ACCESSION AX961914
VERSION AX961914.1 GI:40881324
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 125 18-DEC-2003;
Aashi Kasei Kabushiki Kaisha (JP)
FEATURES
source 1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAF06466.1"
/db_xref="GI:40881325"
/translating="MPGAGSGPSAGDGRRLRLALVLLGWVSASAPSSVPSSTTS
PAFLASGSAQPPPAECPCACESEARTVKVNRNLLLEVPADLPPTVRLNLTGNQ
MTVLPAGAFARQPPPLADEALNLSGNHLKEVCAGAFHLGLRLDLSHNPLTNSAF
VFAGSNASVAPSPLLEILNHI VPPEDORQNGFEGMVAPEGMVAALRSGALRGL
TRLEASNHFELFLPDLIAQLPSLYLDLRNNSLSVLTYSFRLTHLESLELDNAL
KVLHNSLTAEQGLAHVKVFLDNNPWCDCYMDMVAWLKETEVPDKARLTCAPEK
MNRGLDLNSSLDDCDVAVLPQSLOTYSYVFLGIVLALIGAIFLLVLYLNRRKGIKKMMH
NIRDACRDMEGHYRYEINADPRLTNLSSNDV"

ORIGIN
Alignment Scores: 21.2 Length: 2557
Pred. No.: 46.00 Matches: 9
Score: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-22 (1-9) x AX961914 (1-2557)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 1237 TTCTTTCTGCTCGGACTTACTA 1263

RESULT 8
LOCUS MMU012160 7942 bp DNA linear ROD 15-APR-2005
DEFINITION Mus musculus 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS King,K.W., Sheppard,P.C., Westwater,C., Stern,P.L. and Myers,K.A.
TITLE Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10366710
REFERENCE 2 (bases 1 to 7942)
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK
FEATURES
source Location/Qualifiers
1..7942
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="Lambda Dash"
3108..3113
/bound_moiety="Spi"
3114..3119
/bound_moiety="Spi"
3124..5779
/gene="5T4"
3124..3151
/gene="5T4"
3152..3450
/gene="5T4"
3451..5779
/gene="5T4"
3779..5059
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA0931.1"
/db_xref="GI:3805949"
/db_xref="GOA:Q9Z0L0"
/db_xref="InterPro:IPR000372"
misc_binding
misc_binding
gene
exon
intron
exon
CDS

```

```

VFAGSNASVAPSPLLEILNHI VPPEDORQNGFEGMVAPEGMVAALRSGALRGL
TRLEASNHFELFLPDLIAQLPSLYLDLRNNSLSVLTYSFRLTHLESLELDNAL
KVLHNSLTAEQGLAHVKVFLDNNPWCDCYMDMVAWLKETEVPDKARLTCAPEK
MNRGLDLNSSLDDCDVAVLPQSLOTYSYVFLGIVLALIGAIFLLVLYLNRRKGIKKMMH
NIRDACRDMEGHYRYEINADPRLTNLSSNDV"

ORIGIN
Alignment Scores: 21.2 Length: 2557
Pred. No.: 46.00 Matches: 9
Score: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-22 (1-9) x AX961914 (1-2557)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 1237 TTCTTTCTGCTCGGACTTACTA 1263

RESULT 8
LOCUS MMU012160 7942 bp DNA linear ROD 15-APR-2005
DEFINITION Mus musculus 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS King,K.W., Sheppard,P.C., Westwater,C., Stern,P.L. and Myers,K.A.
TITLE Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10366710
REFERENCE 2 (bases 1 to 7942)
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK
FEATURES
source Location/Qualifiers
1..7942
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="Lambda Dash"
3108..3113
/bound_moiety="Spi"
3114..3119
/bound_moiety="Spi"
3124..5779
/gene="5T4"
3124..3151
/gene="5T4"
3152..3450
/gene="5T4"
3451..5779
/gene="5T4"
3779..5059
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA0931.1"
/db_xref="GI:3805949"
/db_xref="GOA:Q9Z0L0"
/db_xref="InterPro:IPR000372"
misc_binding
misc_binding
gene
exon
intron
exon
CDS

```

```
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProtKB/TrEMBL:Q9Z0L0"
/translation="MPGAGSRGPSAGDGLRLRLALVLLGWVSASAPSSSVSPSSSTS
PADFLASGAPQPAERCPAAACESEAAATVKCNRLLEVPADLPYPVYNLFTGNQ
MTVLPAFAFAPQPLADLEALNLSCNHLKEVCAGAFELGLRLDLGHNPLTNLSAF
VFAGSNASVAPSPLLELILNHI VPPEDQRQNGSEGMVAFEGMAAALRSLAUGL
TRLASNHFLPLRDLIAQLPSRLYLRLNNLSLYASYFRNLTHLESJLHLENAL
KVLNLSLAENQGLAHVRFLDNNPWVCCDYMADVMWLKETEVPDVKARLTCAPPEK
MRNGLLDLNSDLDQDAVLPSQTSYVFLGIVLALIGAFLLVLYLNRGIGKKWMH
NIRDACRDMGSEHYRYEINADPLRLNLSNSDV"
3779..3865
/gene="5T4"
3866..5056
/gene="5T4"
/product="ST4 oncofetal trophoblast glycoprotein"
5713..5718
/gene="5T4"
5759..5764
/gene="5T4"

sig_peptide
mat_peptide

polyA_signal
polyA_signal

ORIGIN
Alignment Scores:
Pred. No.: 47 Length: 7942
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-22 (1-9) x MMU012160 (1-7942)

Qy 1 PheLeuPheLeuProArgAspLeuLeu 9
|||||
Db 4460 TTTCTTTTCCTGCTCGGACTTACTA 4486

RESULT 9
HS467D16/c
LOCUS
DEFINITION
HS467D16 143583 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone RP3-467D16 on chromosome 6p22.3-24.1
Contains the 5' end of the SCAL gene for spinocerebellar ataxia 1
(olivopontocerebellar ataxia 1, autosomal dominant, ataxin 1) with
a poly-glutamine (CAG repeat) polymorphism and the 3' part of the
GMPR gene for GMP reductase, Guanosine 5'-monophosphate
oxidoreductase, complete sequence.
AL009031
ACCESSION
AL009031.1 GI:2808422
VERSION
HTG; ataxia; ataxin; GMP reductase; GMPR; Guanosine; monophosphate;
olivopontocerebellar; oxidoreductase; SCAL; spinocerebellar.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 143583)
Tubby.B.
Direct Submission
AUTHORS
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegasanger.ac.uk
JOURNAL
Clone requests: clonerequest@sanger.ac.uk
On Jan 25, 1998 this sequence version replaced gi:2632073.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SwissProt; Tr, TrEMBL; Wp, WORMPEP; Information
on http://www.sanger.ac.uk/projects/Celegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP3-467D16 is from the library RPCI-3 constructed by the group of
```

```
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1..143583
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="R2PD:RPCIP704D16467"
/db_xref="taxon:9606"
/chromosome="6"
/map="p22.3-24.1"
/clone_lib="RPCI-3"
/clone="RP3-467D16"
/join(complement(AL137003.12:78411..86634),
AL031120.1:17022..152296,
complement(AL034375.23:32790..86336),71405..100765)
/gene="SCAL"
/locus_tag="RP3-467D16.1-001"
/join(71565..73481,93016..93546)
/gene="SCAL"
/locus_tag="RP3-467D16.1-001"
/standard_name="OTTHUMP0000016065"
/note="match: proteins: P54253 Q63540"
/codon_start=1
/product="spinocerebellar ataxia 1 (olivopontocerebellar
ataxia 1, autosomal dominant, ataxin 1)"
/protein_id="CAA15622.1"
/db_xref="GI:2808423"
/db_xref="GOA:P54253"
/db_xref="HGNC:10548"
/db_xref="UniProtKB/Swiss-Prot:P54253"
/translation="MKSQERNECLPPKKREIIPATSRSEEEKAPTLPSONHVRVEGTA
WLPFGGRGHGGRHGPAGTSVELGQIGLHKALSTGSDYSPSAPRSVPVATTL
PAAVATPQPTSPVQVYAHLPHTFQFIGSSOYSGTVASPIPSQLIPPTANPVTSAVA
SAAGATPQRSQLEAYSTLLANMGSLSOTPGHKAQEQEQEQEQEQEQEQEQEQEQEQ
CQQQQQHLSRAPGLITPGSPPPAQNQYVHISSPQNTGRTASPAIPVHLHPQTM
IPHTITLGPSPQVWQYADSGSHFVPREATKKAESRLQAIQAKEVINGEMEKSRRY
GAPSDADLGLKAGKSGSPHYPYESRHHVHPSPSDYSRDPGSRASVMVLPNSNTPA
ADLEVQQAATHREASPTLNDKGLHKPKHRSYALSPHTVIQTHSASBPLPVGLPA
TAFYAGTPPVYIGLSGQQAITYAGSLPQHLVIPGTOPLLIIPVGSMDMEASGAAPAI
VTSSPQFAVPHFTVTTALPKSENFENPEALVTQAAVPMVQAQIHLIPVQSVASPARA
PTLPPYFMKSGIILQANGELKKVEDLKTEDFQSAEISNDLKIDSTVERIEDSHSP
GVQVQFVAGHRAQSVLEVYEPFFVFGQWSGCCPERTSQLFDLPCKLSVGDVC
ISLTLNKLNGSVKKGQPDVPASVLLHKSADGLAGSRHRYAQENGINGQSAQMLSE
NGLKFPKMWGLPAAFPFLTKIBSPKPAATKRRWSAPESRKLEKSEDEPPLTLPKPSL
IPEVKICTEGRSNVCK"
Join(complement(AL137003.12:86476..86634),
complement(AL137003.12:79221..79269),
complement(AL137003.12:78411..78525),
AL031120.1:17022..17147,AL031120.1:89016..89143,
AL031120.1:152235..152296,
complement(AL034375.23:86198..86336),
complement(AL034375.23:33172..33322),
complement(AL034375.23:32790..32959),71405..73481,
93016..100765)
/gene="SCAL"
/locus_tag="RP3-467D16.1-001"
/product="spinocerebellar ataxia 1 (olivopontocerebellar
ataxia 1, autosomal dominant, ataxin 1)"
/note="match: ESTs: AA79236 AA909468 AI052569 AI052586
```

Genomic sequence from Human 6

Unpublished
2 (bases 1 to 148750)
Hawkins, T.L., Reeve, M.P., Christoffersen, A., Birren, B.W., Fasman, K.H., Lander, E.S., McKernan, K., Munro, C., Richardson, P., Barna, N., Chang, A., Cooke, P., Daly, M.J., Devon, K., Dewar, K., J., Forrest, C., Gage, D., Geraghty, K., Guitau, G., Hagos, B., Huang, J., Jacotot, L., Lane, M., Lee, K., Mackenzie, J., Marquis, N., McDermott, J., Molla, M., Moloney, N., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Olotu, A., Peterson, K., Rollins, G., Spencer, J., Stilwell, J., Stone, C., Strickland, C., Sydney, K., Traish, A., Wilmer, F., Zemtseva, I. and Zody, M.

Direct Submission
Submitted (15-JUL-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
3 (bases 1 to 148750)
Hawkins, T.L., Reeve, M.P., Birren, B.W., Fasman, K.H., Lander, E.S., McKernan, K., Munro, C., Richardson, P., Barna, N., Chang, A., Christoffersen, A., Cooke, P., Daly, M.J., Devon, K., Dewar, K., Forrest, C., Gage, D., Geraghty, K., Guitau, G., Hagos, B., Huang, J., Jacotot, L., Kirby, A., Lane, M., Mackenzie, J., Marquis, N., McDermott, J., Molla, M., Moloney, N., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Olotu, A., Peterson, K., Rollins, G., Spencer, J., Stilwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L., Traish, A., Wilmer, F., Zemtseva, I. and Zody, M.

Direct Submission
Submitted (31-JUL-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
On Jul 31, 1997 this sequence version replaced gi:2286044.
The Staden databases, finishing information, and all chromatographic files used in the assembly of this clone are available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

FEATURES
source
1. 148750
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="SGII"
/clone_lib="Research Genetics/Cal Tech CIT978SK-B (plates 1-194)"
repeat_region
complement(1035..1336)
/rpt_family="AluSx"
repeat_region
complement(1714..1779)
/rpt_family="LINE2"
repeat_region
1797..1951
/rpt_family="FRAM"
repeat_region
complement(3652..4238)
/rpt_family="LTR10C"
repeat_region
4382..4423
/rpt_family="AT rich"
repeat_region
complement(5988..6047)
/rpt_family="(CA)n"
repeat_region
7059..7347
/rpt_family="POLY A"
repeat_region
8511..8869
/rpt_family="TBE1B"
repeat_region
complement(9175..9377)
/rpt_family="AluJb"
repeat_region
complement(10633..10720)
/rpt_family="MER5A"
repeat_region
complement(10925..11078)
/rpt_family="MER20"
repeat_region
complement(12108..12218)
/rpt_family="(TA)n"

ORIGIN
Alignment Scores:
Pred. No.: 362 Length: 143583
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-22 (1-9) x H8467D16 (1-143583)

QY 1 PheLeuPheLeuProArgAspLeuLeu 9
|||||TATTTTGGCCAGACCTTTG 130435

Db 130461 TTTTATTTTGGCCAGACCTTTG 130435

RESULT 10
AC002326/c
LOCUS
DEFINITION
AC002326
VERSION
AC002326.1 GI:2288970
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 148750)
Hawkins, T.L., Reeve, M.P., Birren, B.W., Fasman, K.H. and Lander, E.S.

REFERENCE
AUTHORS

```
repeat_region 12687. 12750
/rpt family="MIR"
repeat_region 13172. 13433
/rpt family="AluSx"
repeat_region 13450. 13611
/rpt family="MER63B"
repeat_region 13614. 14244
/rpt family="MER7B"
repeat_region 14069. 15308
/rpt family="FLAM_C"
repeat_region 14408. 14660
/rpt family="MER7B"
repeat_region 14968. 14968
/rpt family="AluSc"
repeat_region 14969. 15308
/rpt family="MER7B"
repeat_region 15312. 15411
/rpt family="MER63B"
repeat_region 15670. 15670
/rpt family="AluSp"
repeat_region 15673. 15730
/rpt family="MER63B"
repeat_region 16118. 16283
/rpt family="L1MA6"
repeat_region 16333. 16333
/rpt family="AT rich"
repeat_region 16968. 16968
/rpt family="AluSg"
repeat_region 17180. 17442
/rpt family="AluJo"
repeat_region 17443. 17481
/rpt family="AT rich"
repeat_region 17737. 17877
/rpt family="MERSA"
repeat_region 17909. 18205
/rpt family="AluSg"
repeat_region 19192. 19370
/rpt family="MERSB"
repeat_region 19435. 19472
/rpt family="AT rich"
repeat_region 19523. 19825
/rpt family="AluSx"
repeat_region 19986. 20036
/rpt family="CA)n"
repeat_region 21348. 21649
/rpt family="AluSg"
repeat_region 21781. 21812
/rpt family="AT rich"
repeat_region 22678. 22872
/rpt family="MIR"
repeat_region 23592. 23697
/rpt family="LINE2"
repeat_region 23733. 23866
/rpt family="AluJb"
repeat_region 25793. 26168
/rpt family="MSTA"
repeat_region 27863. 27978
/rpt family="MER30"
repeat_region 28081. 28129
/rpt family="AT rich"
repeat_region 28223. 28436
/rpt family="AluJo"
repeat_region 28499. 28642
/rpt family="AluJo"
repeat_region 29156. 29371
/rpt family="MLT1C"
repeat_region 29409. 29694
/rpt family="AluSg"
repeat_region 29695. 29878
/rpt family="MLT1C"
repeat_region 30397. 30442
/rpt family="LINE2"
repeat_region 30765. 30814
```

```
repeat_region 31035. 31123
/rpt family="(GAAA)n"
repeat_region 31291. 31811
/rpt family="LIPB3"
repeat_region 31846. 32144
/rpt family="AluJb"
repeat_region 32227. 32384
/rpt family="MIR"
repeat_region 33427. 33668
/rpt family="MLT1A2"
repeat_region 36830. 37191
/rpt family="THE1C"
repeat_region 37717. 37748
/rpt family="(CAAA)n"
repeat_region 37750. 38052
/rpt family="AluJb"
repeat_region 38309. 38411
/rpt family="(GAAA)n"
repeat_region 38806. 39127
/rpt family="AluJo"
repeat_region 39356. 39391
/rpt family="AT rich"
repeat_region 40008. 40033
/rpt family="AT rich"
repeat_region 42494. 42754
/rpt family="AluSx"
repeat_region 43281. 43318
/rpt family="(CA)n"
repeat_region 43318. 43353
/rpt family="(GA)n"
repeat_region 43880. 44181
/rpt family="AluSx"
repeat_region 44713. 44831
/rpt family="(TAGA)n"
repeat_region 45202. 45504
/rpt family="AluSx"
repeat_region 45540. 45609
/rpt family="LIPAS"
repeat_region 46081. 46372
/rpt family="AluY"
repeat_region 47354. 47468
```

Alignment Scores:

Pred. No.:	371	Length:	148750
Score:	46.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-774-176-22 (1-9) x AC002326 (1-148750)

QY 1 PheLeuPheLeuProArgAspLeuLeu 9
|||||
Db 140715 TTTTATTTTTCAGACCTTTTG 140689

RESULT 11

AC158516/c

LOCUS

DEFINITION

sequence.

AC158516 AC117768

VERSION

AC158516.2 GI:63025421

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 167046)

ADDITIONAL INFORMATION

AUTHORS

The sequence of Mus musculus BAC clone RP24-511A23

AC158516 167046 bp DNA linear ROD 21-JUN-2005
Mus musculus BAC clone RP24-511A23 from chromosome 9, complete


```

JOURNAL      Unpublished (2001)
REFERENCE    2 (bases 1 to 167046)
AUTHORS      Wilson,R.K.
TITLE        Direct Submission
JOURNAL      Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE    3 (bases 1 to 167046)
AUTHORS      Wilson,R.K.
TITLE        Direct Submission
JOURNAL      Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park
              Parkway, St. Louis, MO 63108, USA
REFERENCE    4 (bases 1 to 167046)
AUTHORS      Wilson,R.K.
TITLE        Direct Submission
JOURNAL      Submitted (21-JUN-2005) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
COMMENT      On May 4, 2005 this sequence version replaced gi:16156412.
              ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
              ----- Summary Statistics
              Center project name: M_BB0511A23
              Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e. phred quality
>30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone, fosmid clone or direct clone walk sequence.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to obtain the consensus sequence. The
assembly was confirmed by restriction digest.
This finishing standard has slightly changed from the previous
Human standard. Specifically, standards for regions of low sequence
complexity (such as dinucleotide repeats and small unit tandem
repeats) have been relaxed. These regions are very prevalent in the
mouse genome, and the return on extended finishing efforts is
minimal.
If a sequence meets the criteria of the above statement, it needs
no comments or tags. If the criteria are not met, such as ambiguous
bases, then the region is duly annotated.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The BAC Library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

This sequence is the entire insert of the clone.

FEATURES             Location/Qualifiers
     source           1..167046
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /clone="RP24-511A23"
                     /clone_lib="RPCI-24"
                     /date="16685..16712"
misc_feature         16685..16712
                     /note="Sequence derived from PCR product of genomic DNA"
unresolved           31565..31779
                     /note="Unresolved simple sequence repeat."

```

```

unresolved         46721..46808
                     /note="Unresolved simple sequence repeat."
unresolved         142336..142347
                     /note="Sequence derived from one plasmid subclone."
ORIGIN
Alignment Scores:
  Pred. No.:      403      Length:      167046
  Score:          46.00    Matches:      9
  Percent Similarity: 100.0%  Conservative: 0
  Best Local Similarity: 100.0%  Mismatches: 0
  Query Match:      100.0%  Indels:      0
  DB:               6      Gaps:      0

US-10-774-176-22 (1-9) x AC158516 (1-167046)
QY      1 PheLeuPheLeuProArgAspLeuLeu 9
Db      110157 TTTCCTTTTCTGCTCGGGACTTACTA 110131

RESULT 12
LOCUS    AX467373              1260 bp      DNA      linear      PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION   AX467373.1 GI:21900603
KEYWORDS . Felis sp.
SOURCE   Felis sp.
ORGANISM Felis sp.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
          Felinae; Felis.
REFERENCE 1
AUTHORS   Myers,K., Drury,N. and Carroll,M.
TITLE     Polypeptide
JOURNAL   Patent: WO 0238612-A 3 16-MAY-2002;
          Oxford Biomedica (UK) Limited (GB)
FEATURES   Location/Qualifiers
            source           1..1260
                                /organism="Felis sp."
                                /mol_type="unassigned DNA"
                                /db_xref="taxon:9687"
ORIGIN
Alignment Scores:
  Pred. No.:      53.5      Length:      1260
  Score:          43.00    Matches:      8
  Percent Similarity: 100.0%  Conservative: 1
  Best Local Similarity: 88.9%  Mismatches: 0
  Query Match:      93.5%  Indels:      0
  DB:               2      Gaps:      0

US-10-774-176-22 (1-9) x AX467373 (1-1260)
QY      1 PheLeuPheLeuProArgAspLeuLeu 9
Db      661 TTCCTCTTCTGCTCGGGACGTACTG 687

RESULT 13
LOCUS    AX821533              1260 bp      DNA      linear      PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION   AX821533.1 GI:39724929
KEYWORDS . Felis catus (cat)
SOURCE   Felis catus
ORGANISM Felis catus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
          Felinae; Felis.
REFERENCE 1
AUTHORS   Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
TITLE     MHC class I peptide epitopes from the human St4 tumor-associated

```

```
antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
  source
    1..1260
    /organism="Felis catus"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 53.5 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.5% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-22 (1-9) x AX821533 (1-1260)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 661 TTCCTCTTCTTCCTCGGACGTACTG 687
RESULT 14
AX821548 AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS Felis catus (cat)
SOURCE
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
    Felinae; Felis.
REFERENCE 1
AUTHORS Carroll,M.O., Harrop,R.O. and Kingman,S.O.
TITLE MHC class II peptide epitope of 54 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
  source
    1..1260
    /organism="Felis catus"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 53.5 Length: 1260
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.5% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-22 (1-9) x AX821548 (1-1260)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 661 TTCCTCTTCTTCCTCGGACGTACTG 687
RESULT 15
AF063939 AF063939 2333 bp mRNA linear ROD 01-JAN-2000
DEFINITION Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA,
complete cds.
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Murinae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2333)
AUTHORS Ninkina,N.N. and Buchman,V.L.
TITLE Structure and expression of the rat 5T4 gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2333)
AUTHORS Buchman,V.L.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK
FEATURES
  source
    1..2333
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /db_xref="taxon:10116"
    /tissue_type="cerebellum"
    /dev_stage="newborn"
  gene
    1..2333
    /gene="5T4"
  5'UTR
    1..363
    /gene="5T4"
    364..1644
    /gene="5T4"
  CDS
    /codon_start=1
    /product="5T4 oncofetal antigen homolog"
    /protein_id="AAF21770.1"
    /db_xref="GI:6650212"
    /translations="MPCAGSRGSPAGDRLRLARLALVLLGWVSASAPSSSLPSSSTS
    PAAFLASGSAQPPPAECRPAACSEAAATVKCVNRNLLEVPADLPVVRNPLTGNQ
    MTVLPAQAPARQPLADLAVNLNGHLEKVGAGAFELPCLRLDLSHNPLTSLAF
    TFGSNVSPTSPLLEILLHIVPPEDQKQSGFEGVAFEGVAAALRSGLALRGL
    HLELASNHFLYLPRLDQLPLKHLDLRNNLSLVSYASFRNLTHLESJHLEDNAL
    KVLHNSLAEWQGLAHVRFVDNNPWYCDCYMADVMSLKYETVVPDKARLTCAPEK
    MNRGLDLTSSDLDCDATALPQSLQTSYVFLGIVLALIGALFLLVLYLNRRGIKWMH
    NTRDADRHMEGYHYRYEINADPSLTNLSNSDV"
    1645..2333
    /gene="5T4"
    2315..2320
    /gene="5T4"
  polyA_signal
    1045 TTTCTTTACCTGCTCGGACTTATTG 1071
ORIGIN
Alignment Scores:
Pred. No.: 82.5 Length: 2333
Score: 43.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.5% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-22 (1-9) x AF063939 (1-2333)
QY 1 PheLeuPheLeuProArgAspLeuLeu 9
Db 1045 TTTCTTTACCTGCTCGGACTTATTG 1071
Search completed: May 27, 2006, 19:34:53
Job time : 3368.6 secs
```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-21

Perfect score: 40

Sequence: 1 ALIGAIFLL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10774176/runat_26052006_091441_24976/app_query.fasta_1
-DB=Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02h
-USER=US10774176 @CGN 1 1 2389 @runat_26052006_091441_24976 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 8: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *
14: Geneseqn2005s: *
15: Geneseqn2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	246	10 ADK11641	Adk11641 Breast ca
2	40	100.0	475	13 ADU11677	Adul1677 Solid tum
3	40	100.0	901	3 AAA27060	Aaa27060 Canine 5T

4	40	100.0	927	6 ABT07721	Abt07721 Breast ca
5	40	100.0	927	8 ABX76333	Abx76333 Lung canc
6	40	100.0	927	10 ADB80503	Adb80503 Ovarian c
7	40	100.0	927	11 ADN38723	Adn38723 Cancer/an
8	40	100.0	973	8 AAD56198	Aad56198 Human LRR
9	40	100.0	1156	6 ABV99349	Abv99349 Human NOV
10	40	100.0	1260	6 ABK87175	Abk87175 cDNA enco
11	40	100.0	1260	10 ADB97513	Adb97513 Feline 5T
12	40	100.0	1260	10 ADB97452	Adb97452 DNA enco
13	40	100.0	1263	3 AAA27058	Aaa27058 Human 5T4
14	40	100.0	1263	4 AAF89736	Aaf89736 Nucleocid
15	40	100.0	1263	6 ABK87174	Abk87174 cDNA enco
16	40	100.0	1281	3 AAA27059	Aaa27059 Mouse 5T4
17	40	100.0	1331	8 AAD56199	Aad56199 Human LRR
18	40	100.0	2020	10 ADJ56299	Adj56299 Human cdn
19	40	100.0	2053	8 ACC51052	Acc51052 Human bla
20	40	100.0	2053	8 ABX76332	Abx76332 Lung canc
21	40	100.0	2053	8 AAD56197	Aad56197 Human LRR
22	40	100.0	2053	8 AAD56200	Aad56200 Human LRR
23	40	100.0	2053	11 ADN38721	Adn38721 Cancer/an
24	40	100.0	2053	12 ADL06473	Adl06473 Human tum
25	40	100.0	2053	12 ADN03961	Adn03961 Antipeori
26	40	100.0	2053	13 ADR25444	Adr25444 Breast ca
27	40	100.0	2053	13 ACN38510	Acn38510 Tumour-as
28	40	100.0	2053	13 ADV35098	Adv35098 Human cdn
29	40	100.0	2053	14 AED17761	Aed17761 Fibrotic
30	40	100.0	2338	5 AAS87175	Aas87175 DNA enco
31	40	100.0	2359	4 AAK94253	Aak94253 Human ful
32	40	100.0	2359	12 ADL30831	Adl30831 Full leng
33	40	100.0	2361	4 AAK94254	Aak94254 Human ful
34	40	100.0	2361	12 ADI26162	Adi26162 Human cdn
35	40	100.0	2361	12 ADL30833	Adl30833 Full leng
36	40	100.0	2557	12 ADI26160	Adi26160 Human cdn
37	40	100.0	2557	12 ADI26158	Adi26158 Human cdn
38	39	97.5	1494	10 ADF02237	Adf02237 Bacterial
39	38	95.0	433	6 ABN60581	Abn60581 Human can
40	38	95.0	567	6 ABN60925	Abn60925 Human can
41	38	95.0	1101	12 ADO35702	Ado35702 Novel mou
42	37	92.5	951	8 ACA28481	Aca28481 Prokaryot
43	36	90.0	225	6 ABN70062	Abn70062 Streptoco
44	36	90.0	225	6 ABN69887	Abn69887 Streptoco
45	36	90.0	453	5 AAS87174	Aas87174 DNA enco

ALIGNMENTS

RESULT 1
ADK11641
ID ADK11641 standard; DNA; 246 BP.
XX
AC ADK11641;
XX
DT 06-MAY-2004 (first entry)
XX
DE Breast cancer differentially expressed gene product #47.
KW ds; cytostatic; gene therapy; DKFp566l133 activity inhibitor;
KW breast cancer; differential expression.
XX
OS Homo sapiens.
XX
PN WO2003057926-A1.
XX
PD 17-JUL-2003.
XX
PF 08-JAN-2003; 2003WO-US000657.
XX
PR 08-JAN-2002; 2002US-0345637P.
XX
(CHIR) CHIRON CORP.
XX
PI Hansen R;
XX

DR WPI; 2003-577534/54.
XX Inhibiting a cancerous phenotype of a cell, useful for treating breast
PT cancer comprises contacting a cancerous mammalian cell with an agent for
PT inhibition of DXFZp5661133 activity.
XX
XX Claim 30; SEQ ID NO 47; 257pp; English.
XX
XX The invention relates to a method of inhibiting a cancerous phenotype of
CC a cell comprises contacting a cancerous mammalian cell with an agent for
CC inhibition of DXFZp5661133 activity. The methods are useful for treating
CC cancer, e.g. breast cancer. This sequence represents a gene product which
CC is differentially expressed in breast cancer cells. The sequence can be
CC used in the method of the invention.
XX
SQ Sequence 246 BP; 77 A; 49 C; 59 G; 61 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 15.7 Length: 246
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-21 (1-9) x ADK11641 (1-246)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 3 GCCTGATAGCGCTATTTCTCTG 29

RESULT 2
ADU11677
ID ADU11677 standard; DNA; 475 BP.
XX
AC ADU11677;
XX
XX 27-JAN-2005 (first entry)
XX
XX Solid tumour prognosis gene seqid 2116.
XX
XX cytostatic; gene therapy; expression profile; solid tumour;
XX peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
XX Unidentified.
XX
XX WO2004097052-A2.
XX
XX 11-NOV-2004.
XX
XX 29-APR-2004; 2004WO-US013587.
XX
XX 29-APR-2003; 2003US-0456067P.
XX
XX 23-JAN-2004; 2004US-0538246P.
XX
XX (AMHP) WYETH.
XX
XX (STRA/) STRAHS A.
XX
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
XX Immerman F, Dorner AJ;
XX
XX WPI; 2004-804779/79.
XX
XX A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
XX Disclosure; Page: 111pp; English.
XX
XX The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially

CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 33.1 Length: 475
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-21 (1-9) x ADU11677 (1-475)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 369 GCCTGATAGCGCTATTTCTCTG 395

RESULT 3
AAA27060
ID AAA27060 standard; DNA; 901 BP.
XX
AC AAA27060;
XX
XX 22-AUG-2000 (first entry)
XX
XX Canine 5T4 tumour-associated antigen gene.
XX
XX Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX ds.
XX
XX Canis sp.
XX
XX Key Location/Qualifiers
XX CDS 1..858
XX /tag= a /product= "5T4 antigen"
XX /tag= b
XX misc_feature 61..74
XX /note= "given in the specification but does not seem to
XX be part of the coding sequence and does not encode any
XX corresponding amino acids"
XX misc_feature 135..146
XX /tag= c

FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 207..216
 FT FT /tag= d
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 277..290
 FT FT /tag= e
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 351..361
 FT FT /tag= f
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 422..436
 FT FT /tag= g
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 497..511
 FT FT /tag= h
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 572..583
 FT FT /tag= i
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 644..653
 FT FT /tag= j
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 714..723
 FT FT /tag= k
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"
 784..801
 FT FT /tag= l
 FT FT /note= "given in the specification but does not seem to
 be part of the coding sequence and does not encode any
 corresponding amino acids"

WO200029428-A2.

25-MAY-2000.

18-NOV-1999; 99WO-GB003859.

18-NOV-1998; 98GB-00025303.

27-JAN-1999; 99GB-00001739.

30-JUL-1999; 99GB-00017995.

(OXFO-) OXFORD BIOMEDICA UK LTD.

Carroll MW, Myers KA;

WPI; 2000-387735/33.

P-PSDB; AAY94351.

FT FT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 response useful in vaccinating against and in treating tumors.

Disclosure; Page 78-79; 79pp; English.

XX XX The present sequence encodes the canine 5T4 tumour-associated antigen
 (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in

CC CC carcinomas but has a highly restricted expression pattern in normal adult
 CC CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC CC induced were inoculated with a virus expression vector containing the
 CC CC human or murine 5T4 gene sequence. The 5T4 antigen was shown to be
 CC CC effective at eliciting an immunotherapeutic anti-tumour response. Both
 CC CC the nucleic acid encoding the antigen and the antigen itself can be used
 CC CC to elicit an immune response, preferably CTL or an antibody response in a
 CC CC subject

XX SQ Sequence 901 BP; 178 A; 246 C; 212 G; 153 T; 0 U; 112 Other;

Alignment Scores: Pred. No.: 68.6 Length: 901

Score: 40.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 3 Gaps: 0

US-10-774-176-21 (1-9) x AAA27060 (1-901)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9

DB 657 GCCTGATAGGCCCATCTTCTACTG 683

RESULT 4

ABT07721

ID ABT07721 standard; DNA; 927 BP.

XX AC ABT07721;

XX DT 14-NOV-2002 (first entry)

XX DE Breast cancer-associated gene sequence 29.

XX Gene; ds; breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.

XX OS Unidentified.

XX PN WO200259377-A2.

XX PD 01-AUG-2002.

XX PF 24-JAN-2002; 2002WO-US002242..

XX PR 24-JAN-2001; 2001US-0263965P.

XX PR 02-FEB-2001; 2001US-0265928P.

XX PR 09-APR-2001; 2001US-00829472.

XX PR 09-APR-2001; 2001US-0282698P.

XX PR 04-MAY-2001; 2001US-0288590P.

XX PR 29-MAY-2001; 2001US-0294443P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC, Afar D;

XX WPI; 2002-583738/62.

XX DR N-PSDB; ABJ05564.

XX Detecting a breast cancer-associated transcript in a patient's cell,
 PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.

XX Claim 9; Page 372; 414pp; English.

XX The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridises to one of the 69 breast cancer-associated gene sequences shown

CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 70.9 Length: 927
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-21 (1-9) x ABT07721 (1-927)

Qy 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
 Db 748 GCCCTGATAGCGCTATTTCTCTCG 774

RESULT 5

ABX76333

ID ABX76333 standard; DNA; 927 BP.

XX AC ABX76333;

DT 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #197.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 XX interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX W0200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

XX 10-MAY-2001; 2001US-0290492P.

XX 09-NOV-2001; 2001US-0339245P.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0334370P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX P-PSDB; ABUS6604.

XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX Claim 22; Page 336; 453pp; English.

XX

CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 70.9 Length: 927
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-21 (1-9) x ABX76333 (1-927)

Qy 1 AlaLeuIleGlyAlaIlePheLeuLeu 9

Db 748 GCCCTGATAGCGCTATTTCTCTCG 774

RESULT 6

ADB80503

ID ADB80503 standard; DNA; 927 BP.

XX AC ADB80503;

XX 04-DEC-2003 (first entry)

XX Ovarian cancer-associated transcript #34.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; ds; gene.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..927

XX /*tag= a

XX W02002102235-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

XX 27-AUG-2001; 2001US-0315287P.

XX 05-SEP-2001; 2001US-0317544P.

XX 13-NOV-2001; 2001US-0350666P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

XX

XX

XX

XX

XX

XX

XX

DR P-PSDB; ADB0504.

XX Detecting an ovarian cancer-associated transcript in a cell from a

PT patient, comprises contacting a biological sample from the patient with a

PT polynucleotide that hybridizes to an ovarian cancer gene.

XX

PS Claim 10; Page 297; 33pp; English.

XX

CC The invention relates to a method of detecting an ovarian cancer-

CC associated transcript in a cell from a patient, by contacting a

CC biological sample from the patient with a polynucleotide that selectively

CC hybridizes to a sequence at least 80% identical to any of one of 80

CC nucleic acid sequences given in the specification. The method is useful

CC in diagnosing ovarian cancer and in identifying and using agents and/or

CC targets that inhibit ovarian cancer. The nucleic acid molecule,

CC polypeptide and the antibody may also be used in detecting ovarian

CC cancers, monitoring and early detection of relapse following treatment,

CC monitoring response to therapy, selecting patients for post-operative

CC chemotherapy or radiation therapy, in selecting mode of therapy,

CC determining tumour prognosis, early detection of pre-cancerous lesions,

CC and as vaccines. This sequence corresponds to one of the nucleic acids

CC used for the detection method of the invention.

XX

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	70.9	Length:	927
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-774-176-21 (1-9) x ADB0503 (1-927)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9

Db 748 GCCCTGATAGGCGCTATTTCTCTCTG 774

RESULT 7

ADN38723

ID ADN38723 standard; cDNA; 927 BP.

XX

AC ADN38723;

XX

DT 17-JUN-2004 (first entry)

XX

DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX

KW Human; differential expression; cancer; angiogenic disorder;

KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;

KW inflammatory disease; autoimmune disease.

KW retinal neovascularisation syndrome; scarring; uterine fibroid;

KW detection; diagnosis; prognosis; drug screening; drug targeting;

KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;

KW vulnery; gene therapy; vaccine; gene; ss.

XX

OS Homo sapiens.

XX

WO2003042661-A2.

XX

22-MAY-2003.

XX

13-NOV-2002; 2002WO-US036810.

XX

13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.

PR 10-JAN-2002; 2002US-0347349P.

PR 08-FEB-2002; 2002US-0355250P.

PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-0368099P.

PR 04-APR-2002; 2002US-0370110P.

PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.

PR 16-JUL-2002; 2002US-0396839P.

PR 22-JUL-2002; 2002US-0397775P.

PR 22-JUL-2002; 2002US-0397845P.

PR 09-SEP-2002; 2002US-0409450P.

XX

PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX

PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX

DR WPI; 2003-468649/44.

DR P-PSDB; ADN38724.

XX

PT Determining the presence or absence of a pathological cell in a patient,

PT useful for diagnosing, prognosing or treating cancer, comprises detecting

PT a nucleic acid in a biological sample.

XX

PS Claim 8; SEQ ID NO 41; 1385pp; English.

XX

CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)

CC whose expression is upregulated or downregulated in specific cancers or

CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a

CC patient by detecting a nucleic acid at least 80% identical to those of

CC the invention or by detecting a polypeptide of the invention. The

CC invention also relates to expression vectors and host cells comprising a

CC nucleic acid of the invention; antibodies which specifically bind a

CC polypeptide of the invention; use of such antibodies for drug targeting;

CC and methods of screening for modulators of activity or expression of the

CC polypeptides and nucleic acids. The nucleic acids, polypeptides,

CC antibodies and methods are useful for diagnosing, prognosing and treating

CC cancer and other conditions such as psoriasis, ischaemia, heart disease,

CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal

CC neovascularisation syndromes, scarring and uterine fibroids. They may

CC also be useful in wound healing and in contraception. The present

CC sequence represents a nucleic acid sequence of the invention.

XX

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	70.9	Length:	927
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-774-176-21 (1-9) x ADN38723 (1-927)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9

Db 748 GCCCTGATAGGCGCTATTTCTCTCTG 774

RESULT 8

AAD56198

ID AAD56198 standard; DNA; 973 BP.

XX

AC AAD56198;

XX

DT 07-AUG-2003 (first entry)

XX

DE Human LRRCAPS related DNA #5.

XX

KW Human; p53 pathway; Leucine rich repeat capricious related protein;

KW LRRCAPS; cancer; gene therapy; ds.

XX

OS Homo sapiens.

```

XX PN WO2003035831-A2.
XX PD 01-MAY-2003.
XX PF 21-OCT-2002; 2002WO-US033540.
XX PR 22-OCT-2001; 2001US-0338733P.
XX PR 15-FEB-2002; 2002US-0357600P.
XX PR 01-MAR-2002; 2002US-0361196P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX DR WPI; 2003-421410/39.
XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
XX PT comprises contacting an assay system comprising a purified leucine rich
XX PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX PS Example 5; Page 74-75; 99pp; English.
XX CC The invention relates to a method of identifying a candidate p53 pathway
XX CC modulating agent. The method involves contacting an assay system
XX CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX CC polypeptide or nucleic acid or its fragment with a test agent and
XX CC detecting a test agent-biased activity, where a difference between the
XX CC test agent-biased activity and the reference activity identifies the test
XX CC agent as a candidate p53 pathway modulating agent. The method is useful
XX CC for identifying a candidate p53 pathway-modulating agent for preparing a
XX CC composition for diagnosing or treating cancer. The invention is useful in
XX CC gene therapy. The present sequence is human LRRCAPS related DNA
XX SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 74.9 Length: 973
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-21 (1-9) x AAD56198 (1-973)
QY 1 AlaLeulleGlyAlaallePheLeuLeu 9
Db 763 GCCCTGATAGCGCTATTTCCTCTG 789

RESULT 9
ABV99349
ID ABV99349 standard; DNA; 1156 BP.
XX AC ABV99349;
XX AC ABV99349;
XX DT 27-JAN-2003 (first entry)
XX DE Human NOV8a coding sequence.
XX KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
XX KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
XX KW neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
XX KW antiinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
XX KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
XX KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
XX KW immune disorder; haematopoietic disorder; cardiovascular disorder;
XX KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
XX KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
XX KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX OS Homo sapiens.
XX

```

(CURA-) CURAGEN CORP.

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 PI Boldog FL, Li L, Zernhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 PI Zhong M;
 XX
 DR WPI; 2002-732824/79.
 DR P-PSDB; ABP70071.

XX New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 PT disorders, and asthma.

XX Claim 16; Page 114-115; 619pp; English.

XX The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious diseases, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, haematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods

XX SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 91.1 Length: 1156
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-21 (1-9) x ABV99349 (1-1156)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9
 DB 979 GCCCTGATAGGCGCTATTTCCTCTG 1005

RESULT 10
 ABK87175
 ID ABK87175 standard; cDNA; 1260 BP.

XX ABK87175;

XX 07-OCT-2002 (first entry)

XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.

XX Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.

XX Felis sp.

XX Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a

FT /product= "5T4 protein"

XX WO200238612-A2.

XX 16-MAY-2002.

XX 13-NOV-2001; 2001WO-GB005004.

XX 13-NOV-2000; 2000WO-GB004317.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Myers K, Drury N, Carroll M;

XX WPI; 2002-557449/59.

XX P-PSDB; AAU98694.

XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.

XX Claim 4; Page 68; 68pp; English.

XX The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus (es).
 CC The present sequence encodes feline 5T4 protein

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 101 Length: 1260
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-21 (1-9) x ABK87175 (1-1260)

Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9

DB 1087 GCCCTGATAGGCGCTATTTCCTCTG 1113

RESULT 11

ADB97513

ID ADB97513 standard; DNA; 1260 BP.

XX ADB97513;

XX 04-DEC-2003 (first entry)

XX Feline 5T4 antigen DNA.

XX Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.

XX Unidentified.

XX

```

FH Key Location/Qualifiers
FT CDS 1..1260
FT /tag= a
FT /product= "Feline 5T4 antigen protein"
XX
XX WO2003068816-A1.
XX
XX 21-AUG-2003.
XX
XX 13-FEB-2003; 2003WO-GB000670.
XX
XX 13-FEB-2002; 2002GB-00003419.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll M, Kingsman S, Redchenko I;
XX
XX WPI; 2003-637141/60.
XX
XX P-PSDB; ADB97520.
XX
XX New major histocompatibility complex class I peptide epitopes from human
XX 5T4 tumor-associated antigen, useful for preventing and/or treating a
XX disease, particularly cancer.
XX
XX Disclosure; Page 67; 73pp; English.
XX
XX The invention relates to a novel Major Histocompatibility Complex (MHC)
XX class I peptide epitope of the 5T4 antigen. The invention further
XX provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
XX sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
XX epitope; a vector system capable of delivering the 5T4 epitope nucleic
XX acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
XX 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
XX comprising the above; a method for treating and/or preventing a disease
XX in a subject by administering the vaccine; an agent capable of binding
XX specifically to the 5T4 epitope and/or its encoding nucleic acid; a method
XX comprising detecting the presence of the 5T4 epitope or its encoding
XX nucleic acid in a subject; and a T cell line or clone capable of
XX specifically recognising the 5T4 epitope in conjunction with an MHC class
XX I molecule. The 5T4 epitope has cytostatic activity. The vaccine
XX comprising the 5T4 epitope or its encoding nucleic acid and the vector
XX system or cell is useful in the prevention and/or treatment of a disease,
XX particularly cancer. The detection method is useful for diagnosing or
XX monitoring the progression of a cancerous disease, and for detecting the
XX presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
XX is useful in the manufacture of a medicament for treating and/or
XX preventing a disease. This polynucleotide sequence represents the feline
XX 5T4 antigen coding DNA of the invention.
XX
XX Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
XX
Alignment Scores:
Pred. No.: 101 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-21 (1-9) x ADB97513 (1-1260)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 1087 GCCCTGATAGTGCCATTTTCTTACTG 1113
RESULT 12
ADB97452
ID ADB97452 standard; DNA; 1260 BP.
XX
XX ADB97452;
XX
XX 04-DEC-2003 (first entry)
XX
XX

```

```

DE DNA encoding feline 5T4 protein.
XX
XX gene; ds; feline; Major Histocompatibility Complex class II; MHC;
XX epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
XX CDS 1..1260
XX /tag= a
XX /product= "Feline 5T4 antigen protein"
XX
XX WO2003068815-A2.
XX
XX 21-AUG-2003.
XX
XX 13-FEB-2003; 2003WO-GB000618.
XX
XX 13-FEB-2002; 2002GB-00003420.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll M, Harrop R, Kingsman S;
XX
XX WPI; 2003-663795/62.
XX
XX P-PSDB; ADB97455.
XX
XX New Major Histocompatibility Complex class II peptide epitope of 5T4,
XX useful for manufacturing a medicament for diagnosing, preventing and/or
XX treating a disease, e.g. cancer.
XX
XX Disclosure; Page 49; 63pp; English.
XX
XX The invention relates to a Major Histocompatibility Complex (MHC) class
XX II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
XX clone has a cytostatic activity, as it is useful in manufacturing a
XX medicament for cytostatic activity, and/or treating a disease, particularly cancer.
XX The methods are useful for detecting T-cells capable of specifically
XX recognising a peptide epitope in conjunction with an MHC molecule, for
XX diagnosing or monitoring the progression of a cancerous disease, or for
XX detecting the presence of a peptide or nucleic acid using an agent. The
XX MHC class II peptide epitope of the invention can be used in gene therapy
XX or as part of a vaccine. This polynucleotide sequence represents the DNA
XX coding for the feline 5T4 protein.
XX
XX Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
XX
Alignment Scores:
Pred. No.: 101 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0
US-10-774-176-21 (1-9) x ADB97452 (1-1260)
QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
Db 1087 GCCCTGATAGTGCCATTTTCTTACTG 1113
RESULT 13
AAA27058
ID AAA27058 standard; DNA; 1263 BP.
XX
XX AAA27058;
XX
XX 22-AUG-2000 (first entry)
XX
XX Human 5T4 tumour-associated antigen gene.
XX
XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX

```

KW ds.
 XX Homo sapiens.
 XX WO200029428-A2.
 XX PD 25-MAY-2000.
 XX PF 18-NOV-1999; 99WO-GB003859.
 XX PR 18-NOV-1998; 98GB-00025303.
 XX PR 27-JAN-1999; 99GB-00001739.
 XX PR 30-JUL-1999; 99GB-00017995.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Carroll MW, Myers KA;
 XX WPI: 2000-387735/33.
 XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte response useful in vaccinating against and in treating tumors.
 XX Example 2; Page 78; 79pp; English.
 XX The present sequence encodes the human 5T4 tumour-associated antigen (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in carcinomas but has a highly restricted expression pattern in normal adult tissues. It appears to be strongly correlated to metastasis in colorectal and gastric cancer. 5T4 antigen may therefore be useful in tumour diagnosis, targeting and immunotherapy. Mice in which tumours had been induced were inoculated with a virus expression vector containing the present sequence. The 5T4 antigen was shown to be effective at eliciting an immunotherapeutic anti-tumour response. Both the nucleic acid encoding the antigen and the antigen itself can be used to elicit an immune response, preferably CTL or an antibody response in a subject
 XX Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 101 Length: 1263
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-21 (1-9) x AAA27058 (1-1263)
 Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9
 |||||
 Db 1090 GCCCTGATAGGCGCTATTTTCTCTCTG 1116
 RESULT 14
 AAF89736
 ID AAF89736 standard; DNA; 1263 BP.
 XX AAF89736;
 XX 23-JUL-2001 (first entry)
 XX Nucleotide sequence of canine 5T4 protein.
 KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
 KW hypersensitivity; autoimmune disease; central nervous system disorder;
 KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
 KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
 KW Helicobacter-related disease; immune disorder; ss.
 XX Canis sp.
 OS Key Location/Qualifiers
 FH 1. .1263
 FT CDS

FT /*tag= a
 XX /product= "5T4"
 XX WO200136486-A2.
 XX PD 25-MAY-2001.
 XX PF 13-NOV-2000; 2000WO-GB004317.
 XX PR 18-NOV-1999; 99WO-GB003859.
 XX PR 15-FEB-2000; 2000GB-00003527.
 XX PR 02-MAR-2000; 2000GB-00005071.
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
 XX Myers KA;
 XX WPI: 2001-343805/36.
 XX P-PSDB; AAB83839.
 XX Use of single chain antibody capable of recognizing a disease associated molecule for manufacturing a medicament for preventing and/or treating a disease condition associated with disease associated molecule.
 XX Disclosure; Fig 26; 118pp; English.
 XX The specification describes the use of a single chain antibody (ScFv), which is capable of recognizing a disease associated molecule in the manufacture of a medicament for the prevention and treatment of a disease condition. The ScFv antibody is useful in the manufacture of a medicament, for affecting a disease in vivo, for preparing a pharmaceutical composition, for in vivo imaging and/or for adjuvant treatment of a disease. The ScFv antibody is also useful for treating inflammatory diseases including arthritis, hypersensitivity, autoimmune diseases, cancers, central nervous system disorders including Parkinson's disease, periodontal diseases, cardiopulmonary diseases, cardiovascular diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-related diseases, and other immune disorders. The present sequence encodes a 5T4 protein, which is used to produce ScFv of the invention
 XX Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 101 Length: 1263
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0
 US-10-774-176-21 (1-9) x AAF89736 (1-1263)
 Qy 1 AlaLeuileGlyAlaIlePheLeuLeu 9
 |||||
 Db 1090 GCCCTGATAGGCGCCATCTTCTCTCTG 1116
 RESULT 15
 ABR87174
 ID ABR87174 standard; cDNA; 1263 BP.
 XX ABR87174;
 XX 07-OCT-2002 (first entry)
 XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
 KW Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX Canis sp.

```
XX Key Location/Qualifiers
FH CDS 1..1263
FT /*tag= a
FT /product= "5T4 protein"
XX
XX WO200238612-A2.
XX
XX 16-MAY-2002.
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Myers K, Drury N, Carroll M;
XX
XX WPI; 2002-557449/59.
XX
XX P-PSDB; AAU98693.
XX
XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
XX polypeptide, useful in preparation of vaccine for treating and/or
XX preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 1; Page 67; 68pp; English.
XX
XX The present invention relates to the isolation of canine and feline
XX oncofoetal leucine-rich glycoproteins known as 5T4, and the
XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
XX a significant proportion of tumours. The sequences of the invention are
XX useful in a pharmaceutical composition for the prevention and/or
XX treatment of tumours or other diseases associated with cell
XX proliferation, infections, and inflammatory conditions in animals,
XX preferably dogs or cats. The compositions may also be used for cancer
XX immunotherapy in these animals. The sequences of the invention may also
XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
XX measurement and localisation of 5T4 in extracts of plasma, urine,
XX tissues, and in cell culture media. Antibodies specific for the 5T4
XX protein are useful for isolating foetal cells from maternal blood. The
XX isolation process may form part of a diagnostic method e.g. the foetal
XX cells may then be subject to biochemical or genetic sampling used for
XX testing foetal abnormalities, or to determine the sex of the foetus(es).
XX The present sequence encodes canine 5T4 protein
XX
XX Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 101 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
```

US-10-774-176-21 (1-9) x ABK87174 (1-1263)

```
QY 1 AlaLeuIleGlyAlaIlePheLeuIleu 9
| | | | | | | | | | | | | | | | | |
Db 1090 GCCGTGATAGCGGCATCTCTCTACTG 1116
```

Search completed: May 27, 2006, 10:38:34
Job time : 378.5 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds

(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-21

Perfect score: 40

Sequence: 1 ALGAIPELL 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSREB.spool/US10774176/runat_26052006_091443_24987/app_query.fasta.1
-DB=GenEmbl -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs04
-USER=US10774176 @CGN 1 1 7524 @runat_26052006_091443_24987 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_scs:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	290	2	CQ687716 Sequence
2	40	100.0	475	2	CQ920916 Sequence
3	40	100.0	901	2	BD249733 Polypepti

SUMMARIES

source

Location/Qualifiers
1. .290
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0.598 Length: 290
Score: 40.00 Matches: 9

4	40	100.0	901	2	AX025013	Sequence
5	40	100.0	901	2	AX316088	Sequence
6	40	100.0	927	2	AX829164	Sequence
7	40	100.0	1156	2	DD161112	Novel Ant
8	40	100.0	1260	2	AX467373	Sequence
9	40	100.0	1260	2	AX821533	Sequence
10	40	100.0	1260	2	AX821548	Sequence
11	40	100.0	1263	2	BD249731	Polypepti
12	40	100.0	1263	2	AX025011	Sequence
13	40	100.0	1263	2	AX149553	Sequence
14	40	100.0	1263	2	AX316086	Sequence
15	40	100.0	1263	2	AX467371	Sequence
16	40	100.0	1281	2	BD249732	Polypepti
17	40	100.0	1281	2	AX025012	Sequence
18	40	100.0	1281	2	AX316087	Sequence
19	40	100.0	2053	2	CQ731678	Sequence
20	40	100.0	2053	2	DD174290	METHODS O
21	40	100.0	2053	5	HS5T40A	Homo sapien
22	40	100.0	2333	6	AF063939	Rattus no
23	40	100.0	2359	2	BD127282	Primer fo
24	40	100.0	2359	2	CQ782724	Sequence
25	40	100.0	2359	5	AK074786	Homo sapi
26	40	100.0	2361	2	AX961916	Sequence
27	40	100.0	2361	2	BD127283	Primer fo
28	40	100.0	2361	2	CQ782726	Sequence
29	40	100.0	2361	5	AK074790	Homo sapi
30	40	100.0	2361	6	BC087011	Rattus no
31	40	100.0	2379	5	BC037161	Homo sapi
32	40	100.0	2423	6	BC058198	Mus muscu
33	40	100.0	2557	2	AX961912	Sequence
34	40	100.0	2557	2	AX961914	Sequence
35	40	100.0	2714	5	AB168308	Macaca fa
36	40	100.0	5551	5	HS012159	Homo sapi
37	40	100.0	7942	6	MMU012160	Mus muscu
38	40	100.0	121909	5	HSJ492P14	Human DNA
39	40	100.0	167046	6	AC158516	Mus muscu
40	40	100.0	210237	12	AC128294	Rattus no
41	40	100.0	239076	12	AC106962	Rattus no
42	40	100.0	243801	12	AC172361	Bos tauru
43	39	97.5	1494	2	AR377516	Sequence
44	39	97.5	158710	6	AC126539	Mus muscu
45	39	97.5	206384	6	AC158916	Mus muscu

ALIGNMENTS

RESULT 1
CQ687716
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Liew, C.C., Marshall, W.E. and Zhang, H.
Compositions and methods relating to osteoarthritis
Patent: WO 02070737-A 32642 12-SEP-2002;
Chondrogene Inc. (CA)
Location/Qualifiers
1. .290
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CQ687716
Sequence 32642 from Patent WO02070737.
CQ687716
CQ687716.1 GI:42218962
linear PAT 03-FEB-2004


```

DB:
US-10-774-176-21 (1-9) x AX025013 (1-901)
Qy 1 AlaLeuIIeGlyAlaIlePheLeuLeu 9
Db 657 GCCCTGATAGGCGCCATCTTCTACTG 683

RESULT 5
AX316088
LOCUS AX316088 901 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 3 from Patent EP1160323.
ACCESSION AX316088
VERSION AX316088.1 GI:17899280
KEYWORDS
SOURCE
ORGANISM
Canis sp.
Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE
1 Carroll,M.W. and Myers,K.A.
AUTHORS
TITLE
5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..901
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN
Alignment Scores:
Pred. No.: 2.07 Length: 901
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-21 (1-9) x AX316088 (1-901)
Qy 1 AlaLeuIIeGlyAlaIlePheLeuLeu 9
Db 657 GCCCTGATAGGCGCCATCTTCTACTG 683

RESULT 6
AX829164
LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1 Mack,D.H., Gish,K.C. and Afar,D.
AUTHORS
TITLE
Methods of diagnosis of breast cancer, compositions and methods of
screening for modulators of breast cancer
JOURNAL
Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)
FEATURES
source
1..927
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.13 Length: 927

```

```

Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-21 (1-9) x AX829164 (1-927)
Qy 1 AlaLeuIIeGlyAlaIlePheLeuLeu 9
Db 748 GCCCTGATAGGCGCTATTTTCCTCTG 774

RESULT 7
DD161112
LOCUS DD161112 1156 bp DNA linear PAT 23-NOV-2005
DEFINITION Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
Encoding The Antigens, and Methods of Use.
ACCESSION DD161112
VERSION DD161112.1 GI:83967439
KEYWORDS JP 2005508604-A/23.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1 (bases 1 to 1156)
AUTHORS Padigaru,M., Shenoy,S.G., Pochart,P.F., Kekuda,R., Gusev,V.Y.,
Zhong,M., Jr,R.J.T., Casman,S.J., Li,L., Miller,C.E.,
Patturajan,M., Anderson,D.W., Malyankar,U.M., Voss,E.Z.,
Spaderna,S.K., Gorman,L., Spytek,K.A., Liu,X., Burgess,C.E.,
Pena,C.E.A., Gerlach,V., Smithson,G., Mezes,P.D., Rastelli,L.,
Boldog,F.L., Guo,X., Vernet,C.A.M., Gangolli,E.A., Tchernev,V.T.
and Zerhusen,B.D.
TITLE Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
Encoding The Antigens, and Methods of Use
JOURNAL Patent: JP 2005508604-A 23 07-APR-2005;
Muraidhara Padigaru,Suresh Shenoy,Remesh Kekuda,Vladimir Gusev,
Pascalle Pochart,Mei Zhong,Luca Rastelli,Peter Mezes, Glennad
Smithson, Xiaojia Guo,Valerie Gerlach,Stacie Casman,Ferend
Boldog,Li Li, Bryan Zerhusen,Velizar Tchernev,Esha Gangolli, Corine
Vernet, Carol Pena, Catherine Burgess,Xiaohong Liu,Kimberly
Spytek,Linda Gorman, Steven Spaderna,Edward Voss,Uriel
Malyankar,David Anderson, Meera Patturajan,Charles Miller,Raymond J
Taupier Jr
COMMENT
OS Homo sapiens
PN JP 2005508604-A/23
PD 07-APR-2005
PF 08-MAR-2002 JP 2002571827
PR 19-JUN-2001 US 60/299310,18-JUN-2001 US 60/299027, PR
31-MAY-2001 US 60/294889,31-MAY-2001 US 60/294899, PR
30-MAY-2001 US 60/294485,09-MAR-2001 US 60/274849, PR
13-MAR-2001 US 60/275579,03-MAY-2001 US 60/288342, PR
02-MAY-2001 US 60/288066,30-APR-2001 US 60/287424, PR
13-APR-2001 US 60/283675,04-APR-2001 US 60/281194, PR
02-APR-2001 US 60/280802,02-APR-2001 US 60/280822, PR
03-MAY-2001 US 60/288528,15-MAY-2001 US 60/291190, PR
16-MAY-2001 US 60/291240,16-MAY-2001 US 60/291099, PR
12-MAR-2001 US 60/275235,08-MAR-2001 US 60/274101, PR
08-MAR-2001 US 60/274281,08-MAR-2001 US 60/274322, PR
08-MAR-2001 US 60/274194,02-APR-2001 US 60/280900, PR
30-MAR-2001 US 60/280233,30-MAR-2001 US 60/279995, PR
20-MAR-2001 US 60/277327,20-MAR-2001 US 60/277338, PR
19-MAR-2001 US 60/276994,16-MAR-2001 US 60/276776, PR
14-MAR-2001 US 60/276000,13-MAR-2001 US 60/275601, PR
13-MAR-2001 US 60/275578,20-MAR-2001 US 60/277239, PR
20-MAR-2001 US 60/277321,21-MAR-2001 US 60/277791, PR
22-MAR-2001 US 60/277833,23-MAR-2001 US 60/278152, PR
26-MAR-2001 US 60/278894,27-MAR-2001 US 60/278999, PR
27-MAR-2001 US 60/279036,28-MAR-2001 US 60/279344, PR
19-JUN-2001 US 60/299303,10-JUL-2001 US 60/304354, PR
31-JUL-2001 US 60/309198,03-DEC-2001 US 60/337426, PR
03-DEC-2001 US 60/338092,21-NOV-2001 US 60/332094, PR
14-NOV-2001 US 60/333272,14-NOV-2001 US 60/332271, PR

```

```

14-NOV-2001 US 60/332272,14-NOV-2001 US 60/332172, PR
14-NOV-2001 US 60/333184,31-OCT-2001 US 60/335301, PR
18-OCT-2001 US 60/330380,27-SEP-2001 US 60/325430, PR
27-SEP-2001 US 60/325681,12-SEP-2001 US 60/318770, PR
10-SEP-2001 US 60/318462,03-JAN-2002 US 60/345705, PR
04-DEC-2001 US 60/337185,08-MAR-2002 US 10/093463, PR
16-AUG-2001 US 60/312903
PI muralidhara padigar,suresh g shenoy,pascale f-g pochart, PI
ramesh kekuda,
PI vladimir y gusev,mei zhong,raymond j taupier jr,stacie j PI
casman,li li,
PI charles e miller,meera patturajan,david w anderson,uriel m PI
PI malyankar,
PI edward z voss,steven k spaderna,linda gorman,kimberly PI a
PI xiaohong liu,catherine e burgess,carol e a pena,valerie PI
gerlach,
PI glenda smithson,peter d mezes,luca rastelli,ferenc l boldog,
PI xiaojia guo,
PI corine a m vernet,esha a gangolli,velizar t tchernev,bryan d
PI zerhusen
CC
FH Key Location/Qualifiers
FT CDS (1)..(1150).
FEATURES
source
1..1156
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.71 Length: 1156
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-21 (1-9) x DD161112 (1-1156)
QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 979 GCCCTGATAGCGCTATTTCTCTG 1005
RESULT 8
AX467373
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
Felis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 Myers,K., Drury,N. and Carroll,M.
POLYPEPTIDE
Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
Location/Qualifiers
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"
ORIGIN
Alignment Scores:
Pred. No.: 2.98 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

```

```

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-21 (1-9) x AX467373 (1-1260)
QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1087 GCCCTGATAGGTGCCATTTTCTTACTG 1113
RESULT 9
AX821533
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
AUTHORS
TITLE MHC class I peptide epitopes from the human St4 tumor-associated antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
Location/Qualifiers
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 2.98 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-21 (1-9) x AX821533 (1-1260)
QY 1 AlaLeulleGlyAlaIlePheLeuLeu 9
Db 1087 GCCCTGATAGGTGCCATTTTCTTACTG 1113
RESULT 10
AX821548
LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE
ORGANISM
Felis catus (cat)
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
Felinae; Felis.
REFERENCE
1 Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
AUTHORS
TITLE MHC class II peptide epitope of St4 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1260
Location/Qualifiers
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN

```



```
Alignment Scores:
Pred. No.: 2.98 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-21 (1-9) x AX821548 (1-1260)

Qy 1 AlaLeuilleGlyAlaIlePheLeuLeu 9
|||||
Db 1087 GCCCTGATAGGCGCATTTCTCTACTG 1113

RESULT 11
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1263)
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PR 18-NOV-1999 JP 2000582415
PF 30-JUL-1999 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
PI MILES WILLIAM CARROLL,KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K14/065,
PC C07K19/00,
PC C12N15/00
CC Polypeptide
FH Key Location/Qualifiers
FT source
FT 1..1263 /organism="Homo sapiens (human)".
FEATURES
source
1..1263 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.99 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-21 (1-9) x BD249731 (1-1263)

Qy 1 AlaLeuilleGlyAlaIlePheLeuLeu 9
|||||
Db 1090 GCCCTGATAGGCGCATTTCTCTCTG 1116

RESULT 12
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS

SOURCE Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo sapiens
REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source
1..1263 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.99 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-21 (1-9) x AX149553 (1-1263)

Qy 1 AlaLeuilleGlyAlaIlePheLeuLeu 9
|||||
Db 1090 GCCCTGATAGGCGCATTTCTCTACTG 1116

RESULT 14
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE 1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
Ellard,F.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1263 Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="5T4"
```

Search completed: May 27, 2006, 19:35:24
Job time : 3359.6 secs

DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE S4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2.99 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-21 (1-9) x AX316086 (1-1263)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
|||||
Db 1090 GCCCTGATAGCGCATTTCTCTCTG 1116

RESULT 15

AX467371 1263 bp DNA linear PAT 16-JUL-2002
LOCUS AX467371
DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS
SOURCE Canis sp.
ORGANISM Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE 1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1..1263
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN

Alignment Scores:
Pred. No.: 2.99 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-21 (1-9) x AX467371 (1-1263)

QY 1 AlaLeuIleGlyAlaIlePheLeuLeu 9
|||||
Db 1090 GCCCTGATAGCGCATTTCTCTCTG 1116

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-20

Perfect score: 46

Sequence: 1 FLTGQNTV 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB/spool/US1074176/runat_26052006_091441_24976/app_query.fasta.1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02h
-USER=US1074176 @CGN 1 1 2389 @runat_26052006_091441_24976 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAFN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 8.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*
15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	1281	3	AAA27059 Mouse 5T4
2	46	100.0	2557	12	ADI26160 Human cDN
3	46	100.0	2557	12	ADI26158 Human cDN

C	4	46	100.0	2557	12	ADO35939
	5	42	91.3	335913	5	AAI61371
	6	42	91.3	335913	5	AAI61372
	7	39	84.8	140036	6	AAS98600
	8	38	82.6	343	8	ABX40605
	9	38	82.6	505	14	ACL56146
	10	38	82.6	1260	6	ABK87175
	11	38	82.6	1260	10	ADB97513
	12	38	82.6	1260	10	ADB97452
	13	38	82.6	1263	3	AAA27058
	14	38	82.6	1263	4	AAF89736
	15	38	82.6	1263	6	ABK87174
	16	38	82.6	1331	8	AAD56199
	17	38	82.6	2020	10	ADJ56299
	18	38	82.6	2053	8	ACC51052
	19	38	82.6	2053	8	ABX76332
	20	38	82.6	2053	8	AAD56197
	21	38	82.6	2053	8	AAD56200
	22	38	82.6	2053	11	ADN38721
	23	38	82.6	2053	12	ADL06473
	24	38	82.6	2053	12	ADN03961
	25	38	82.6	2053	13	ADR25444
	26	38	82.6	2053	13	ACN38510
	27	38	82.6	2053	13	ADV35098
	28	38	82.6	2338	5	AAS87175
	29	38	82.6	2359	4	AAK94253
	30	38	82.6	2359	12	ADL30831
	31	38	82.6	2359	12	ADL30831
	32	38	82.6	2361	4	AAK94254
	33	38	82.6	2361	12	ADI26162
	34	38	82.6	2361	12	ADL30833
C	35	36	78.3	201	13	ADS41219
	36	36	78.3	317	3	AAA79371
	37	36	78.3	458	4	AAL02538
	38	36	78.3	563	12	ADO41135
C	39	36	78.3	654	5	AAS94238
C	40	36	78.3	735	4	AAL07288
C	41	36	78.3	792	6	ABQ42606
	42	36	78.3	792	6	ABQ42607
C	43	36	78.3	795	6	ABQ42584
	44	36	78.3	795	6	ABQ42585
45	36	78.3	984	8	ACA35557	

ALIGNMENTS

RESULT 1

AAA27059

ID AAA27059 standard; DNA; 1281 BP.

XX

AC AAA27059;

XX

DT 22-AUG-2000 (first entry)

XX

DE Mouse 5T4 tumour-associated antigen gene.

XX

XX Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;

KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;

KW ds.

XX

OS Mus musculus.

XX

XX WO200029428-A2.

XX

PD 25-MAY-2000.

XX

PF 18-NOV-1999; 99WO-GB003859.

XX

PR 18-NOV-1998; 98GB-00025303.

PR

PR 27-JAN-1999; 99GB-00001739.

PR

PR 30-JUL-1999; 99GB-00017995.

XX

PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll MW, Myers KA;
 XX WPI; 2000-387735/33.
 XX
 PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.
 XX
 PS Example 2; Page 78; 79pp; English.
 XX
 CC The present sequence encodes the mouse 5T4 tumour-associated antigen
 CC ('TAA'). The TAA 5T4 is a glycoprotein which is widely expressed in
 CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The 5T4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject. The
 CC present sequence appears in GenBank at accession number AJ012160
 XX
 SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2.04 Length: 1281
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-20 (1-9) x AAA27059 (1-1281)
 QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
 |||||
 DB 289 TTCCTTACCGCAACACGATGACCGTG 315
 RESULT 2
 ADI26160
 ID ADI26160 standard; cDNA; 2557 BP.
 XX
 AC ADI26160;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #63.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 PF 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00137505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 XX (ASAH) ASAHI KASEI KK.

PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX
 DR WPI; 2004-122214/12.
 DR P-PSDB; ADI26161.
 XX
 PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 PS Claim 4; SEQ ID NO 125; 1368pp; English.
 XX
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.58 Length: 2557
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-20 (1-9) x ADI26160 (1-2557)
 QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
 |||||
 DB 844 TTCCTTACCGCAACACGATGACCGTG 870
 RESULT 3
 ADI26158
 ID ADI26158 standard; cDNA; 2557 BP.
 XX
 AC ADI26158;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #62.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.

XX PF 05-JUN-2003; 2003WO-JP007123.
XX PR 05-JUN-2002; 2002JP-00164257.
XX PR 06-JUN-2002; 2002US-0385912P.
XX PR 26-DEC-2002; 2002JP-00377326.
XX PR 27-DEC-2002; 2002US-0436467P.
XX PR 15-MAY-2003; 2003JP-00137505.
XX PR 16-MAY-2003; 2003US-0470836P.
XX PA (ASAH) ASahi KASEI KK.
XX XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
XX PI WPI; 2004-122214/12.
XX DR P-PSDB; ADI26159.
XX XX
XX PT New signal transducer and activator of transcription 6 activation
XX PT promoting purified protein, for diagnosing and treating disease
XX PT associated with activation/inhibition of transcription factor e.g.
XX PT diabetes and cancer.
XX PS Claim 4; SEQ ID NO 123; 1368pp; English.
XX XX
XX CC The invention relates to a purified protein promoting signal transducer
XX CC and activator of transcription 6 activation (STAT6). The protein is
XX CC useful for the producing an antibody, which involves administering the
XX CC protein or its epitope-bearing fragments to a non-human animal as an
XX CC antigen. The nucleic acid is useful for diagnosing a disease or
XX CC susceptibility to a disease related to expression or activity of the
XX CC protein. A transformant expressing the protein is useful for screening
XX CC compounds which inhibit or promote STAT6 activation. A transformant
XX CC expressing the protein is useful for producing a pharmaceutical
XX CC composition. Compositions, antibodies and antisense molecules are useful
XX CC for the treating a disease associated with STAT6 activation such as
XX CC allergic diseases, inflammation, autoimmune diseases, diabetes,
XX CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
XX CC for treating disease associated with STAT6 activation and/or prevention
XX CC of Th hyperactive diseases. Compositions are also useful in rheumatoid
XX CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
XX CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
XX CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
XX CC activity. The protein or nucleic acid is effectively useful for screening
XX CC compounds for treating and preventing disease associated with excessive
XX CC activation or inhibition of STAT6. The present sequence represents a
XX CC human cDNA encoding a protein which promotes STAT6 activation.
XX SQ Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.58 Length: 2557
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-20 (1-9) x ADI26158 (1-2557)
Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
Db 844 TTCCTTACCGGCACACCATGATGACCGTG 870

RESULT 4
ID ADO35939/C
XX ADO35939 standard; DNA; 2557 BP.
XX AC ADO35939;
XX XX
XX DT 26-AUG-2004 (first entry)
XX XX
XX DE Novel mouse gene sequence #612.
XX XX

KW mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
KW viral disorder; ds; gene.
XX Mus sp.
XX OS
XX PN WO2004046310-A2.
XX XX
XX PD 03-JUN-2004.
XX XX
XX PF 24-OCT-2003; 2003WO-US033948.
XX XX
XX PR 15-NOV-2002; 2002US-0426916P.
XX PR 04-DEC-2002; 2002US-0431158P.
XX PR 05-DEC-2002; 2002US-0431445P.
XX PR 05-DEC-2002; 2002US-0431606P.
XX PR 09-JUN-2003; 2003US-0476621P.
XX PR 09-JUN-2003; 2003US-0476632P.
XX PR 08-JUL-2003; 2003US-0485217P.
XX PR 08-JUL-2003; 2003US-0485359P.
XX PR 08-AUG-2003; 2003US-0493332P.
XX PR 08-AUG-2003; 2003US-0493356P.
XX XX
XX PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX XX
XX PI Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
XX XX WPI; 2004-431966/40.
XX DR
XX PT New mouse nucleic acid molecules and polypeptides, useful for treating
XX PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
XX PT disease or thrombosis.
XX PS Claim 1; SEQ ID NO 612; 263pp; English.
XX XX
XX CC The invention comprises 744 novel mouse DNA sequences (genes). The DNA
XX CC sequences of the invention are useful for treating cancer, psoriasis,
XX CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
XX CC immune disorders, bacterial disorders and viral disorders. The present
XX CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
XX CC present DNA sequence is not shown in the specification, but has been
XX CC retrieved from the WIPO website.
XX SQ Sequence 2557 BP; 610 A; 794 C; 688 G; 465 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.58 Length: 2557
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-20 (1-9) x ADO35939 (1-2557)
Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
Db 57 TTCCTTACCGGCACACCATGATGACCGTG 31

RESULT 5
AAI61371
ID AAI61371 standard; DNA; 335913 BP.
XX XX
XX AC AAI61371;
XX XX
XX DT 16-OCT-2001 (first entry)
XX XX
XX DE Soybean 240017 region G3, SEQ ID NO: 2.
XX XX
XX KW Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;
KW SCN resistance; Rhg1; Rhg4; SCN resistant allele; plant breeding;
KW 240017 region G3; 318013 region A3; 515002 region G2; ds.
XX XX

OS Glycine max.
 XX WO200151627-A2.
 XX 19-JUL-2001.
 XX 05-JAN-2001; 2001WO-US000552.
 XX 07-JAN-2000; 2000US-0174880P.
 XX (MONS) MONSANTO CO.
 XX Hauge BM, Wang ML, Parsons JD, Parnell LD;
 XX WPI; 2001-425872/45.
 XX P-PSDB; AAM42214.
 XX New purified nucleic acid for producing a soybean plant having soybean
 PT cyst nematode resistance and for use in plant breeding programs.
 XX Claim 2; Page 204-400; 1353pp; English.
 XX The invention relates to nucleic acid molecules from regions of the
 CC soybean genome which are associated with soybean cyst nematode (SCN)
 CC resistance. The nucleic acids are used to transform plants, and can
 CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
 CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
 CC of soybean plants and for introgressing SCN resistance or partial SCN
 CC resistance into soybean plants. They can also be used in plant breeding
 CC programmes. The invention also relates to proteins encoded by such
 CC nucleic acid molecules, as well as antibodies capable of recognising
 CC these proteins. The present sequence is a nucleic acid molecule provided
 CC in the specification
 XX SQ Sequence 335913 BP; 114579 A; 53403 C; 53026 G; 114905 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.12e+04 Length: 335913
 Score: 42.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.3% Indels: 0
 DB: 5 Gaps: 0
 US-10-774-176-20 (1-9) x AAI61371 (1-335913)
 QY 1 PheLeuThrGlyAenGlnMetThr 8
 DB 269585 TTCTTAACGGAAATCAGATGACA 269608
 RESULT 6
 AAI61372
 ID AAI61372 standard; DNA; 335913 BP.
 XX
 AC AAI61372;
 XX
 XX 16-OCT-2001 (first entry)
 DT
 XX Soybean 240017 region G3, SEQ ID NO: 3.
 DE
 XX Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;
 KW SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;
 KW 240017 region G3; 318013 region A3; 515002 region G2; ds.
 XX Glycine max.
 OS
 XX WO200151627-A2.
 XX 19-JUL-2001.
 XX 05-JAN-2001; 2001WO-US000552.
 XX 07-JAN-2000; 2000US-0174880P.
 XX (MONS) MONSANTO CO.
 XX Hauge BM, Wang ML, Parsons JD, Parnell LD;
 XX WPI; 2001-425872/45.
 XX P-PSDB; AAM42214.
 XX New purified nucleic acid for producing a soybean plant having soybean
 PT cyst nematode resistance and for use in plant breeding programs.
 XX Claim 2; Page 204-400; 1353pp; English.
 XX The invention relates to nucleic acid molecules from regions of the
 CC soybean genome which are associated with soybean cyst nematode (SCN)
 CC resistance. The nucleic acids are used to transform plants, and can
 CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
 CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
 CC of soybean plants and for introgressing SCN resistance or partial SCN
 CC resistance into soybean plants. They can also be used in plant breeding
 CC programmes. The invention also relates to proteins encoded by such
 CC nucleic acid molecules, as well as antibodies capable of recognising
 CC these proteins. The present sequence is a nucleic acid molecule provided
 CC in the specification
 XX SQ Sequence 335913 BP; 114579 A; 53403 C; 53026 G; 114905 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.12e+04 Length: 335913
 Score: 42.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.3% Indels: 0
 DB: 5 Gaps: 0
 US-10-774-176-20 (1-9) x AAI61371 (1-335913)
 QY 1 PheLeuThrGlyAenGlnMetThr 8
 DB 269585 TTCTTAACGGAAATCAGATGACA 269608
 RESULT 6
 AAI61372
 ID AAI61372 standard; DNA; 335913 BP.
 XX
 AC AAI61372;
 XX
 XX 16-OCT-2001 (first entry)
 DT
 XX Soybean 240017 region G3, SEQ ID NO: 3.
 DE
 XX Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;
 KW SCN resistance; rhg1; Rhg4; SCN resistant allele; plant breeding;
 KW 240017 region G3; 318013 region A3; 515002 region G2; ds.
 XX Glycine max.
 OS
 XX WO200151627-A2.
 XX 19-JUL-2001.
 XX 05-JAN-2001; 2001WO-US000552.
 XX 07-JAN-2000; 2000US-0174880P.

XX (MONS) MONSANTO CO.
 XX Hauge BM, Wang ML, Parsons JD, Parnell LD;
 XX WPI; 2001-425872/45.
 XX P-PSDB; AAM42215.
 XX New purified nucleic acid for producing a soybean plant having soybean
 PT cyst nematode resistance and for use in plant breeding programs.
 XX Claim 2; Page 400-595; 1353pp; English.
 XX The invention relates to nucleic acid molecules from regions of the
 CC soybean genome which are associated with soybean cyst nematode (SCN)
 CC resistance. The nucleic acids are used to transform plants, and can
 CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
 CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
 CC of soybean plants and for introgressing SCN resistance or partial SCN
 CC resistance into soybean plants. They can also be used in plant breeding
 CC programmes. The invention also relates to proteins encoded by such
 CC nucleic acid molecules, as well as antibodies capable of recognising
 CC these proteins. The present sequence is a nucleic acid molecule provided
 CC in the specification
 XX SQ Sequence 335913 BP; 114582 A; 53398 C; 53027 G; 114906 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.12e+04 Length: 335913
 Score: 42.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.3% Indels: 0
 DB: 5 Gaps: 0
 US-10-774-176-20 (1-9) x AAI61372 (1-335913)
 QY 1 PheLeuThrGlyAenGlnMetThr 8
 DB 269585 TTCTTAACGGAAATCAGATGACA 269608
 RESULT 7
 AAS98600
 ID AAS98600 standard; DNA; 140036 BP.
 XX
 AC AAS98600;
 XX
 XX 12-MAR-2002 (first entry)
 DT
 XX Human genomic DNA for PHIP/NDRP.
 DE
 XX PHIP; pleckstrin homology domain-interacting protein; NDRP; ds;
 KW neuronal differentiation-related protein; insulin receptor substrate;
 KW IRS; signal transducer and activator of transcription; STAT;
 KW transgenic animal; diabetes mellitus type 2; hyperglycaemia;
 KW myotonic muscular dystrophy; acanthosis; nigricans; retinopathy;
 KW nephropathy; arteriosclerosis; peripheral arterial disease; cancer;
 KW adenocarcinoma; leukaemia; breast cancer; prostate cancer; colon cancer;
 KW ovarian cancer; autoimmune disease; inflammation; immunodeficiency.
 XX Homo sapiens.
 OS
 XX WO200185785-A2.
 XX 15-NOV-2001.
 XX 10-MAY-2001; 2001WO-CA000673.
 XX 11-MAY-2000; 2000US-0203561P.
 XX (ROZA/) ROZAKIS-ADCOCK M.
 XX (FARH/) FARHANG-FALLAH J.
 XX (CHEN/) CHENG A.

XX Rosakis-Adcock M, Farhang-Fallah J, Cheng A;
 PI WPI; 2002-041586/05.
 XX Novel Pleckstrin homology domain interacting protein recruiting proteins
 PT of insulin receptor substrate family, and signal transducer and activator
 PT of transcription factors to their receptors, useful to treat diabetes.
 XX Disclosure; Page 99-133; 139pp; English.
 XX The invention relates to an isolated Pleckstrin homology domain
 CC interacting protein (PHIP) that recruits proteins of the insulin receptor
 CC substrate (IRS) family, and signal transducer and activator of
 CC transcription (STAT) transcription factors, to receptors that interact
 CC with and phosphorylate the proteins and STAT transcription factors, the
 CC nucleic acid encoding PHIP (nPHIP), a nucleic acid which binds to nPHIP
 CC or regions of it, analogues, fragments or allelic variants of PHIP or
 CC nPHIP, a nucleic acid sequence having substantial sequence identity or
 CC sequence similarity with a nucleic acid sequence fully defined human
 CC neuronal differentiation-related protein (NDRP) nucleic acid sequence or
 CC its exons as given in the specification, expression vectors and host
 CC cells expressing the nucleic acids, anti-PHP antibodies, and a transgenic
 CC animal not already expressing PHIP. The nucleic acids, proteins and
 CC antibodies are useful for diagnosis and treatment of a condition
 CC associated with an insulin receptor (e.g. diabetes mellitus type 2,
 CC hyperglycaemia, myotonic muscular dystrophy, acanthosis, nigricans,
 CC retinopathy, nephropathy, arteriosclerosis, peripheral arterial disease)
 CC or cancer (e.g. adenocarcinoma, leukaemia, breast cancer, prostate
 CC cancer, colon cancer, ovarian cancer and many others given in the
 CC specification), autoimmune disease, inflammation and immunodeficiency.
 CC The protein is also useful for discovering or testing compounds which may
 CC be either enhancers or inhibitors of PHIP function. The present sequence
 CC is genomic DNA encoding PHIP and NDRP
 XX Sequence 140036 BP; 48500 A; 25843 C; 22499 G; 43194 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1.93e+04 Length: 140036
 Score: 39.00 Matches: 7
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 87.5% Mismatches: 0
 Query Match: 84.8% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-20 (1-9) x AAS98600 (1-140036)
 QY 1 PheLeuThrGlyAsnGlnMetThr 8
 Db 18628 TATCTTACTGCTAATCAATGACA 18651
 RESULT 8
 ABX40605
 ID ABX40605 standard; cDNA; 343 BP.
 AC ABX40605;
 XX 20-FEB-2003 (first entry)
 DT 20-FEB-2003 (first entry)
 DE Bovine EST associated with lactation/muscle/fat deposition #5770.
 XX Bovine; ss: EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX Bos Taurus.
 XX US2002137139-A1.
 BN 26-SEP-2002.
 PD 24-SEP-2001; 2001US-00960352.
 XX 24-SEP-2001; 2001US-00960352.
 XX

PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 DR New nucleic acid associated with lactation, and muscle and fat
 XX deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 PT Claim 2; SEQ ID NO 5770; 245pp; English.
 PS The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 1512 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX Sequence 343 BP; 40 A; 146 C; 108 G; 49 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 28.6 Length: 343
 Score: 38.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 82.6% Indels: 0
 DB: 8 Gaps: 0
 US-10-774-176-20 (1-9) x ABX40605 (1-343)
 QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
 Db 179 TTCTTCACGGCAACCACTGCGCGTG 205
 RESULT 9
 ACL56146
 ID ACL56146 standard; cDNA; 505 BP.
 XX ACL56146;
 AC 24-MAR-2005 (first entry)
 DT 24-MAR-2005 (first entry)
 XX Human colon cancer differentially expressed polynucleotide, SEQ ID:2281.
 DE Differential expression; diagnosis; therapy; drug screening; cancer;
 KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
 XX

sg.

Homo sapiens.

WO2005000087-A2.

06-JAN-2005.

13-MAY-2004; 2004WO-US015421.

03-JUN-2003; 2003US-0475872P.

(CHIR) CHIRON CORP.

Randazzo F, Moler E, Escobedo J, Garcia PD;
WPI; 2005-075421/08.

New isolated polynucleotides, which are differentially expressed in colon cancer cell, useful for treating cancer, e.g. colon cancer, breast cancer, or pancreatic cancer.

Claim 1; SEQ ID NO 2281; 97pp; English.

The invention relates to 9672 polynucleotides (AC153866-ACL63537) which are differentially expressed in colon cancer cells. The invention also relates to vectors and host cells comprising a differentially expressed polynucleotide of the invention; a method for detecting a cancerous cell by detection of a gene product of the polynucleotides; a method for inhibiting a cancerous phenotype of a cell by inhibiting a gene product of the polynucleotides; a method of treating an individual with cancer by administration of a modulator of a gene product of the polynucleotides; and an isolated antibody that specifically binds to a polypeptide encoded by one of the 9672 polynucleotides. The polynucleotides, polypeptides, antibodies, and methods are useful for the detection of cancerous cells; for the diagnosis, prognosis and management of cancer; for the identification of agents that modulate the phenotype of cancerous cells; for the identification of therapeutic targets for cancer chemotherapy; and for the treatment of cancer, especially colon cancer and metastasized colon cancer, but also breast or pancreatic cancer. The polynucleotides are also useful as a source of probes or primers for use in diagnostic methods. The differentially expressed polynucleotides or their encoded proteins can additionally be used as vaccines to modulate primary immune responses for the prevention or treatment of cancer. The present sequence represents a specifically claimed polynucleotide which is differentially expressed in colon cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 505 BP; 68 A; 202 C; 149 G; 85 T; 0 U; 1 Other;

Alignment Scores:	45.1	Length:	505
Pred. No.:	38.00	Matches:	7
Score:	88.9%	Conservative:	1
Percent Similarity:	77.8%	Mismatches:	1
Best Local Similarity:	82.8%	Indels:	0
Query Match:	14	Gaps:	0
DB:			

US-10-774-176-20 (1-9) x ACL56146 (1-505)

QY 1 PheIeuthrGlyAenGlnMethrVal 9
|||||:|:|:|

Db 180 TTCCTACCGCAACACGCTGCGGTG 206

RESULT 10

ABK87175

ID ASK87175 standard; cDNA; 1260 BP.

XX AC

XX ABK87175;

DT 07-OCT-2002 (first entry)

```

XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
DE Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
XX cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX
OS Felis sp.
XX
XX Key Location/Qualifiers
FH CDS 1..1260
FT /*tag= a
FT /product= "5T4 protein"
XX
FN WO200238612-A2.
XX
XX 16-MAY-2002.
PD
XX
XX 13-NOV-2001; 2001WO-GB005004.
PF
XX
XX 13-NOV-2000; 2000WO-GB004317.
PR
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX Myers K, Drury N, Carroll M;
PI WPI; 2002-557449/59.
PT P-PSDB; RAU98694.
DR
XX
XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
PS Claim 4; Page 68; 68pp; English.
XX
XX The present invention relates to the isolation of canine and feline
CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 5T4 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 5T4
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus(es).
CC The present sequence encodes feline 5T4 protein
XX
SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 131 Length: 1260
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 82.6% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-20 (1-9) x ABX87175 (1-1260)

Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
||| ||||||| ||||| :||
Db 286 TTCCTCACC CGCATCACTGGCGGTG 312

RESULT 11
ADB97513
ID ADB97513 standard; DNA; 1260 BP.
```


XX AC ADB97513;
 XX DT 04-DEC-2003 (first entry)
 XX DE Feline 5T4 antigen DNA.
 XX KW Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.
 XX OS Unidentified.
 XX FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"
 XX PN WO2003068816-A1.
 XX XX 21-AUG-2003.
 XX PF 13-FEB-2003; 2003WO-GB000670.
 XX PR 13-FEB-2002; 2002GB-00003419.
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PI Carroll M, Kingsman S, Redchenko I;
 XX WPI; 2003-637141/60.
 DR P-PSDB; ADB97520.
 XX XX New major histocompatibility complex class I peptide epitopes from human
 PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.
 XX PS Disclosure; Page 67; 73pp; English.
 CC The invention relates to a novel Major Histocompatibility Complex (MHC)
 CC class I peptide epitope of the 5T4 antigen. The invention further
 CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
 CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
 CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 5T4 epitope and/its encoding nucleic acid; a method
 CC comprising detecting the presence of the 5T4 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 5T4 epitope in conjunction with an MHC class
 CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC particularly cancer. The detection method is useful for diagnosing or
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline
 CC 5T4 antigen coding DNA of the invention.
 XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 131 Length: 1260
 Score: 38.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 82.6% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-20 (1-9) x ADB97513 (1-1260)

QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
 DB 286 TTCTCACCAGCAATCAGCTGCCGTG 312

RESULT 12

ADB97452
 ID ADB97452 standard; DNA; 1260 BP.
 XX AC ADB97452;
 XX DT 04-DEC-2003 (first entry)
 XX DE DNA encoding feline 5T4 protein.
 XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
 KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
 XX OS Unidentified.
 XX FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"
 XX PN WO2003068815-A2.
 XX XX 21-AUG-2003.
 XX PF 13-FEB-2003; 2003WO-GB000618.
 XX PR 13-FEB-2002; 2002GB-00003420.
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX PI Carroll M, Harrop R, Kingsman S;
 XX WPI; 2003-663795/62.
 DR P-PSDB; ADB97455.
 XX XX New Major Histocompatibility Complex class II peptide epitope of 5T4,
 PT useful for manufacturing a medicament for diagnosing, preventing and/or
 PT treating a disease, e.g. cancer.
 XX PS Disclosure; Page 49; 63pp; English.
 CC The invention relates to a Major Histocompatibility Complex (MHC) class
 CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
 CC clone has a cytostatic activity, as it is useful in manufacturing a
 CC medicament for preventing and/or treating a disease, particularly cancer.
 CC The methods are useful for detecting T-cells capable of specifically
 CC recognising a peptide epitope in conjunction with an MHC molecule, for
 CC diagnosing or monitoring the progression of a cancerous disease, or for
 CC detecting the presence of a peptide or nucleic acid using an agent. The
 CC MHC class II peptide epitope of the invention can be used in gene therapy
 CC or as part of a vaccine. This polynucleotide sequence represents the DNA
 CC coding for the feline 5T4 protein.
 XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 131 Length: 1260
 Score: 38.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 82.6% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-20 (1-9) x ADB97452 (1-1260)

QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
 DB 286 TTCTCACCAGCAATCAGCTGCCGTG 312

```

RESULT 13
ID AAA27058 standard; DNA; 1263 BP.
XX AC AAA27058;
XX DT 22-AUG-2000 (first entry)
XX DE Human 5T4 tumour-associated antigen gene.
XX KW Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
XX KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX OS Homo sapiens.
XX PN WO200029428-A2.
XX PD 25-MAY-2000.
XX PF 18-NOV-1999; 99WO-GB003859.
XX PR 18-NOV-1998; 98GB-00025303.
XX PR 27-JAN-1999; 99GB-00001739.
XX PR 30-JUL-1999; 99GB-00017995.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Carroll MW, Myers KA;
XX DR WPI; 2000-387735/33.
XX PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
XX PT response useful in vaccinating against and in treating tumors.
XX PS Example 2; Page 78; 79pp; English.
XX CC The present sequence encodes the human 5T4 tumour-associated antigen
XX CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
XX CC carcinomas but has a highly restricted expression pattern in normal adult
XX CC tissues. It appears to be strongly correlated to metastasis in colorectal
XX CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
XX CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
XX CC induced were inoculated with a virus expression vector containing the
XX CC present sequence. The 5T4 antigen was shown to be effective at eliciting
XX CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
XX CC the antigen and the antigen itself can be used to elicit an immune
XX CC response, preferably CTL or an antibody response in a subject
XX SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 132 Length: 1263
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 82.6% Indels: 0
DB: 3 Gaps: 0

US-10-774-176-20 (1-9) x AAA27058 (1-1263)
QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
DB 289 TTCCTTACCGGCACACGCTGGCGGTG 315

RESULT 14
ID AAF89736 standard; DNA; 1263 BP.
XX AC AAF89736;
XX DT 23-JUL-2001 (first entry)

```

```

XX DE Nucleotide sequence of canine 5T4 protein.
XX KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
XX KW hypersensitivity; autoimmune disease; central nervous system disease;
XX KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
XX KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
XX KW Helicobacter-related disease; immune disorder; ss.
XX OS Canis sp.
XX PH Key Location/Qualifiers
XX FT 1..1263
XX FT /*tag= a
XX FT /product= "5T4"
XX PN WO200136486-A2.
XX PD 25-MAY-2001.
XX PF 13-NOV-2000; 2000WO-GB004317.
XX PR 18-NOV-1999; 99WO-GB003859.
XX PR 15-FEB-2000; 2000GB-00003527.
XX PR 02-MAR-2000; 2000GB-00005071.
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX PI Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM,
XX PI Myers KA;
XX DR WPI; 2001-343805/36.
XX DR P-PSDB; AAB83839.
XX PT Use of single chain antibody capable of recognizing a disease associated
XX PT molecule for manufacturing a medicament for preventing and/or treating a
XX PT disease condition associated with disease associated molecule.
XX PS Disclosure; Fig 26; 118pp; English.
XX CC The specification describes the use of a single chain antibody (ScFv),
XX CC which is capable of recognizing a disease associated molecule in the
XX CC manufacture of a medicament for the prevention and treatment of a disease
XX CC condition. The ScFv antibody is useful in the manufacture of a
XX CC medicament, for affecting a disease in vivo, for preparing a
XX CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
XX CC treatment of a disease. The ScFv antibody is also useful for treating
XX CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
XX CC diseases, cancers, central nervous system disorders including Parkinson's
XX CC disease, periodontal diseases, cardiopulmonary diseases, cardiovascular
XX CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
XX CC related diseases, and other immune disorders. The present sequence
XX CC encodes a 5T4 protein, which is used to produce ScFv of the invention
XX SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 132 Length: 1263
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 82.6% Indels: 0
DB: 4 Gaps: 0

US-10-774-176-20 (1-9) x AAF89736 (1-1263)
QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
DB 289 TTCCTTACCGGCACACGCTGGCGGTG 315

RESULT 15
ID ABK87174 standard; CDNA; 1263 BP.

```

Search completed: May 27, 2006, 10:38:27
Job time : 417.5 secs

```

XX ABK87174;
AC
XX
XX 07-OCT-2002 (first entry)
XX
XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
DE
XX
XX Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
XX foetal abnormality; foetal sex determination; gene; ss.
XX
XX Canis sp.
OS
XX
XX
XX Key Location/Qualifiers
XX 1. .1263
FT CDS /*tag= a
FT /product= "5T4 protein"
FT
XX
XX WO200238612-A2.
PN
XX
XX 16-MAY-2002.
PD
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Myers K, Drury N, Carroll M;
PI
XX
XX WPI; 2002-557449/59.
DR
XX P-PSDB; RAU98693.
XX
XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 1; Page 67; 68pp; English.
XX
XX The present invention relates to the isolation of canine and feline
XX oncofoetal leucine-rich glycoproteins known as 5T4, and the
XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
XX a significant proportion of tumours. The sequences of the invention are
XX useful in a pharmaceutical composition for the prevention and/or
XX treatment of tumours or other diseases associated with cell
XX proliferation, infections, and inflammatory conditions in animals,
XX preferably dogs or cats. The compositions may also be used for cancer
XX immunotherapy in these animals. The sequences of the invention may also
XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
XX measurement and localisation of 5T4 in extracts of plasma, urine,
XX tissues, and in cell culture media. Antibodies specific for the 5T4
XX protein are useful for isolating foetal cells from maternal blood. The
XX isolation process may form part of a diagnostic method e.g. the foetal
XX cells may then be subject to biochemical or genetic sampling used for
XX testing foetal abnormalities, or to determine the sex of the foetus(es).
XX The present sequence encodes canine 5T4 protein
XX
XX Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 132 Length: 1263
Score: 38.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 82.6% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-20 (1-9) x ABK87174 (1-1263)

Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
Db 289 TTCCTCAGCGGCACACCGACTGGCGGTG 315

```

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds

(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-20

Perfect score: 46

Sequence: 1 FLTNQMTV 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB.spool/US10774176/runat_26052006.091443_24987/app_query.fasta_1
-DB=GenEmbl -OPMT=fastap -SUPFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=extc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss04
-USER=US10774176 @CGN 1.1 7524 @runat_26052006.091443_24987 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

- 1: gb env:*
- 2: gb pat:*
- 3: gb ph:*
- 4: gb pl:*
- 5: gb pr:*
- 6: gb ro:*
- 7: gb ste:*
- 8: gb sy:*
- 9: gb un:*
- 10: gb vi:*
- 11: gb ov:*
- 12: gb btg:*
- 13: gb in:*
- 14: gb om:*
- 15: gb ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	1281	2 BD249732	Polypeptide
2	46	100.0	1281	2 AX025012	Sequence
3	46	100.0	1281	2 AX316087	Sequence

4	46	100.0	2333	6 AF063939	Rattus no
5	46	100.0	2361	6 BC087011	Rattus no
6	46	100.0	2423	6 BC058198	Mus muscu
7	46	100.0	2557	2 AX961912	Sequence
8	46	100.0	2557	2 AX961914	Sequence
9	46	100.0	7942	6 MMU012160	Mus muscu
C 10	46	100.0	167046	6 AC158516	Mus muscu
C 11	46	100.0	210237	12 AC128294	Rattus no
C 12	46	100.0	239076	12 AC106962	Rattus no
C 13	43	93.5	189229	12 AC130132	Rattus no
C 14	43	93.5	228414	12 AC128563	Rattus no
C 15	42	91.3	179668	6 AC124399	Mus muscu
C 16	42	91.3	214405	6 AC121871	Mus muscu
C 17	42	91.3	230705	12 AC130996	Rattus no
C 18	42	91.3	250467	12 AC094737	Rattus no
C 19	42	91.3	335913	2 AX196295	Sequence
C 20	42	91.3	335913	2 AX196296	Sequence
C 21	41	89.1	927	10 D89676	Chlorella v
C 22	41	89.1	123576	5 AC015969	Homo sapi
C 23	40	87.0	140962	12 CR387991	Danio rer
C 24	40	87.0	189050	15 AL646077	Ralstonia
C 25	40	87.0	261502	12 AC118131	Rattus no
C 26	39	84.8	578	7 G49021	SHGC-78189
C 27	39	84.8	1119	11 AY874346	Xenopus f
C 28	39	84.8	1135	11 AY874348	Xenopus b
C 29	39	84.8	1141	11 AY874347	Xenopus f
C 30	39	84.8	1141	11 AY874350	Xenopus l
C 31	39	84.8	1142	11 AY874344	Xenopus w
C 32	39	84.8	6598	6 MMLICE03	Mus musculu
C 33	39	84.8	100802	5 AC110759	Homo sapi
C 34	39	84.8	110000	12 AC111775_1	Continuation (2 of
C 35	39	84.8	113116	12 AC177017	Strongylo
C 36	39	84.8	140040	2 AX328485	Sequence
C 37	39	84.8	149948	12 AC068852	Homo sapi
C 38	39	84.8	163204	12 AL356749	Homo sapi
C 39	39	84.8	166054	6 AC134539	Mus muscu
C 40	39	84.8	170294	12 CR854985	Danio rer
C 41	39	84.8	170387	5 AL356776	Human DNA
C 42	39	84.8	173067	12 AC156544	Mus muscu
C 43	39	84.8	176104	12 AC016193	Homo sapi
C 44	39	84.8	176846	12 AC167999	Bos tauru
C 45	39	84.8	184876	12 AC181606	Bos tauru

ALIGNMENTS

RESULT 1	BD249732	1281 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD249732				
DEFINITION	Polypeptide.				
ACCESSION	BD249732				
VERSION	BD249732.1 GI:33059502				
KEYWORDS	JP 2002530060-A/2.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;				
TITLE	Sciurognathi; Muroidea; Muridae; Murinae; Mus.				
JOURNAL	1 (bases 1 to 1281)				
COMMENT	Carroll, M.W. and Myers, K.A.				
	Patent: JP 2002530060-A 2 17-SEP-2002;				
	OXFORD BIOMEDICA LTD				
	OS Mus musculus (mouse)				
	PN JP 2002530060-A/2				
	PD 17-SEP-2002				
	PF 18-NOV-1999 JP 2000582415				
	PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR				
	30-JUL-1999 GB 9917995.4				
	PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS				
	PC C12N15/09;A61K39/00;A61K48/00;A61P35/00;C07K7/06;C07K14/065,				
	PC C07K19/00,				
	PC C12N15/00				

```
CC Polypeptide
FH Key Location/Qualifiers
FT source 1..1281
FT /organism="Mus musculus (mouse)"

FEATURES
source 1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 3.6 Length: 1281
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-20 (1-9) x BD249732 (1-1281)
Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
Db 289 TTCCTTACCGCAACCAAGATGACCGTG 315

RESULT 2
AX025012 1281 bp DNA linear PAT 15-SEP-2000
LOCUS
DEFINITION Sequence 2 from Patent WO0029428.
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 Carroll,M.W. and Myers,K.A.
Polypeptide
TITLE Patent: WO 0029428-A 2 25-MAY-2000;
JOURNAL CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source 1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 3.6 Length: 1281
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-20 (1-9) x AX025012 (1-1281)
Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
Db 289 TTCCTTACCGCAACCAAGATGACCGTG 315

RESULT 3
AX316087 1281 bp DNA linear PAT 14-DEC-2001
LOCUS
DEFINITION Sequence 2 from Patent EP1160323.
ACCESSION AX316087
VERSION AX316087.1 GI:17899279
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 Carroll,M.W. and Myers,K.A.
Polypeptide
TITLE Patent: WO 0029428-A 2 25-MAY-2000;
JOURNAL CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source 1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 3.6 Length: 1281
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-20 (1-9) x AX025012 (1-1281)
Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
Db 289 TTCCTTACCGCAACCAAGATGACCGTG 315

RESULT 4
AX063939 2333 bp mRNA linear ROD 01-JAN-2000
LOCUS
DEFINITION Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA,
complete cds.
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 2333)
Ninkina,N.N. and Buchman,V.L.
Structure and expression of the rat 5T4 gene
Unpublished
2 (bases 1 to 2333)
Buchman,V.L.
Direct Submission
Submitted (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK
FEATURES
source 1..2333
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="newborn"
1..2333
/gene="5T4"
1..363
/gene="5T4"
364..1644
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal antigen homolog"
/protein_id="A0P21770.1"
/db_xref="GI:6650212"
/translation="MPGAGSRQPSAGDGRRLRLARLALLVLLGWVSAPSSSLPSSSTS
PAAFLASGAPPPAECRCPAAECSEAAATKVCNRLLEVPADLPFYRNLFITGNQ
MTVLPAQAPROPPLADLAVLNLSGNHLKEVGAGAFHLPGRLRLDLSINPLTNLSAF
TFAGSNVSVSTPSPLLEILNHI VPPEDQRQNGSPFGMVAFEGMVAALRLRSGLALRGL"

gene
5'UTR
CDS
```

HHLELASNHFLYLPDRLLDQLPSLKHLDLRNNSLVSLTVASPRNLTHLESILHEDNAL
 KVLHNSTLAEQGLAHVRVFLDNNPWCDYMDVMVSLWKETEVPVDPKARLTCAFPPEK
 MNRGLDLTSSDLDCDADTLPSQSLQTSYVFLGIVLALIGAILFLVLYLNRRGKIKKWMH
 NNRDRCRDMGEGHYRIEINADPSLTNLSSNSDV"
 1645..2333
 /gene="5T4"
 2315..2320
 /gene="5T4"

polyA_signal

ORIGIN

Alignment Scores:

Pred. No.: 6.67 Length: 2333
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-20 (1-9) x AF063939 (1-2333)

Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9
 |||||
 Db 652 TTCTCACTGGCAACCAAGATGACCGTG 678

RESULT 5

BC087011

LOCUS

DEFINITION

Rattus norvegicus trophoblast glycoprotein, mRNA (cdna clone

MGC:93332 IMAGE:7193411), complete cds.

ACCESSION

BC087011

VERSION

BC087011.1 GI:56268819

KEYWORDS

MGC.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

1 (bases 1 to 2361)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalls, D.E.,

Schmerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2361)

Director MGC Project.

Direct Submission

Submitted (02-DEC-2004) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapb-remail.nih.gov

Tissue Procurement: Howard Jacobs

cDNA Library Preparation: Express Genomics

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 186 Row: 0 Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

source

1..2361
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="MGC:93332 IMAGE:7193411"
 /tissue type="Heart, rat (Brown Norway)"
 /clone_lib="NIH MGC_234"
 /lab_host="DH10B"
 /note="Vector: pExpress1"
 1..2361
 /gene="tpbg"
 /note="synonym: 5T4"
 /db_xref="GeneID:83684"
 /db_xref="RGD:621453"
 364..1644
 /gene="tpbg"
 /codon_start=1
 /product="Tpbq protein"
 /protein_id="AAH87011.1"
 /db_xref="GI:56268820"
 /db_xref="GeneID:83684"
 /db_xref="RGD:621453"
 /translation="MPGAGSRGSPAGDRLRLARLALVLIGWVSAGAPSSSLSSSTS
 PAATLQSGSAQPPAECPCACESEARTKCNENLLEVPADLPYVSNLFTGNQ
 MTVLPAGAPAPPLADLAVLNSGNHKEVGAGFEHLPLRLDLSHNPILNLSAP
 TPAGNSVSTPSPLELILNHIIVPEDQONGSFQGVAFEGWVAALSKSLGALRGL
 HHLELASNHFLYLPDRLLDQLPSLKHLDLRNNSLVSLTVASPRNLTHLESILHEDNAL
 KVLHNSTLAEQGLAHVRVFLDNNPWCDYMDVMVSLWKETEVPVDPKARLTCAFPPEK
 MNRGLDLTSSDLDCDADTLPSQSLQTSYVFLGIVLALIGAILFLVLYLNRRGKIKKWMH
 NNRDRCRDMGEGHYRIEINADPSLTNLSSNSDV"

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.: 6.76 Length: 2361
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-20 (1-9) x BC087011 (1-2361)

Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9

|||||

Db 652 TTCTCACTGGCAACCAAGATGACCGTG 678

RESULT 6

BC058198

LOCUS

DEFINITION

Mus musculus trophoblast glycoprotein, mRNA (cdna clone MGC:68145

IMAGE:5353871), complete cds.

ACCESSION

BC058198

VERSION

BC058198.1 GI:34849573

KEYWORDS

MGC.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 2423)

BC058198 2423 bp mRNA linear ROD 21-OCT-2003

Mus musculus trophoblast glycoprotein, mRNA (cdna clone MGC:68145

IMAGE:5353871), complete cds.

BC058198

BC058198.1 GI:34849573

MGC.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 2423)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaeetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skaleka, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2423)
Strausberg, R.
Direct Submission
Submitted (15-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigr.nih.gov
Akhter, N., Ayele, K., Beckett, S., Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Trowgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 123 Row: p Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

source
1. .2423
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:68145 IMAGE:5353871"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma 5 month old virgin mouse."
/clone_lib="NCI_GCAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2423
/gene="Tpbp"
/note="synonym: 574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
402. .1682
/gene="Tpbp"

gene

translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSTS
PAAFLASGAQPPPAERCPAACESEAARTVKVNRNLLVLPADLPYVNRNLTGNQ
MTVLPAQAFAPQPLADLEALNSGNHLKEVCAGAFEHLGLRRLDLSHNPITNLGAF
VFAGSNASVAPSPLEELILNHI VPPEDORQNGSPFGMVAFEGMVAALRSLGLRGL
TRLEASNHFLPLRDLLAQLPSRLYLDLRNNSLSVLTYSFNRNLTLSLELDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYCNMADVMVWLKETEVPDKARLTCAFPPEK

CDS

translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSTS
PAAFLASGAQPPPAERCPAACESEAARTVKVNRNLLVLPADLPYVNRNLTGNQ
MTVLPAQAFAPQPLADLEALNSGNHLKEVCAGAFEHLGLRRLDLSHNPITNLGAF
VFAGSNASVAPSPLEELILNHI VPPEDORQNGSPFGMVAFEGMVAALRSLGLRGL
TRLEASNHFLPLRDLLAQLPSRLYLDLRNNSLSVLTYSFNRNLTLSLELDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYCNMADVMVWLKETEVPDKARLTCAFPPEK

/codon_start=1
/product="Tpbp protein"
/protein_id="AAH58198.1"
/db_xref="GI:34849574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
/translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSTS
PAAFLASGAQPPPAERCPAACESEAARTVKVNRNLLVLPADLPYVNRNLTGNQ
MTVLPAQAFAPQPLADLEALNSGNHLKEVCAGAFEHLGLRRLDLSHNPITNLGAF
VFAGSNASVAPSPLEELILNHI VPPEDORQNGSPFGMVAFEGMVAALRSLGLRGL
TRLEASNHFLPLRDLLAQLPSRLYLDLRNNSLSVLTYSFNRNLTLSLELDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYCNMADVMVWLKETEVPDKARLTCAFPPEK
MENRGLDLSNSSLDCDAVLPSQSLQSYVFLGLVLAIGALIFLLVLYLNRRKGIKKWMMH
NIRDCARDHMEGYHYRYEINADPRLTNLSSNSDV"
642. .1262
/gene="Tpbp"
/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [Function unknown]"
/db_xref="CDD:COG4886"
1299. .1415
/gene="Tpbp"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

misc_feature

translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSTS
PAAFLASGAQPPPAERCPAACESEAARTVKVNRNLLVLPADLPYVNRNLTGNQ
MTVLPAQAFAPQPLADLEALNSGNHLKEVCAGAFEHLGLRRLDLSHNPITNLGAF
VFAGSNASVAPSPLEELILNHI VPPEDORQNGSPFGMVAFEGMVAALRSLGLRGL
TRLEASNHFLPLRDLLAQLPSRLYLDLRNNSLSVLTYSFNRNLTLSLELDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYCNMADVMVWLKETEVPDKARLTCAFPPEK
MENRGLDLSNSSLDCDAVLPSQSLQSYVFLGLVLAIGALIFLLVLYLNRRKGIKKWMMH
NIRDCARDHMEGYHYRYEINADPRLTNLSSNSDV"
642. .1262
/gene="Tpbp"
/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [Function unknown]"
/db_xref="CDD:COG4886"
1299. .1415
/gene="Tpbp"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

misc_feature

translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSTS
PAAFLASGAQPPPAERCPAACESEAARTVKVNRNLLVLPADLPYVNRNLTGNQ
MTVLPAQAFAPQPLADLEALNSGNHLKEVCAGAFEHLGLRRLDLSHNPITNLGAF
VFAGSNASVAPSPLEELILNHI VPPEDORQNGSPFGMVAFEGMVAALRSLGLRGL
TRLEASNHFLPLRDLLAQLPSRLYLDLRNNSLSVLTYSFNRNLTLSLELDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYCNMADVMVWLKETEVPDKARLTCAFPPEK
MENRGLDLSNSSLDCDAVLPSQSLQSYVFLGLVLAIGALIFLLVLYLNRRKGIKKWMMH
NIRDCARDHMEGYHYRYEINADPRLTNLSSNSDV"
642. .1262
/gene="Tpbp"
/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [Function unknown]"
/db_xref="CDD:COG4886"
1299. .1415
/gene="Tpbp"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

ORIGIN

Alignment Scores:
Pred. No.: 6.94 Length: 2423
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-20 (1-9) x BC058198 (1-2423)
QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
DB 690 TTCCTTACCGGCACACCATGACCGTG 716

RESULT 7
AX961912
LOCUS
DEFINITION
Sequence 123 from Patent WO03104277.
AX961912
ACCESSION
AX961912.1 GI:40881322
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556. .1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881323"
/translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSTS
PAAFLASGAQPPPAERCPAACESEAARTVKVNRNLLVLPADLPYVNRNLTGNQ
MTVLPAQAFAPQPLADLEALNSGNHLKEVCAGAFEHLGLRRLDLSHNPITNLGAF
VFAGSNASVAPSPLEELILNHI VPPEDORQNGSPFGMVAFEGMVAALRSLGLRGL
TRLEASNHFLPLRDLLAQLPSRLYLDLRNNSLSVLTYSFNRNLTLSLELDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYCNMADVMVWLKETEVPDKARLTCAFPPEK

CDS

translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSTS
PAAFLASGAQPPPAERCPAACESEAARTVKVNRNLLVLPADLPYVNRNLTGNQ
MTVLPAQAFAPQPLADLEALNSGNHLKEVCAGAFEHLGLRRLDLSHNPITNLGAF
VFAGSNASVAPSPLEELILNHI VPPEDORQNGSPFGMVAFEGMVAALRSLGLRGL
TRLEASNHFLPLRDLLAQLPSRLYLDLRNNSLSVLTYSFNRNLTLSLELDNAL
KVLHNSTLAEWQGLAHVKVFLDNNPWVCDYCNMADVMVWLKETEVPDKARLTCAFPPEK

MNRGLDLSNDDCDVLPQSLQTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"

ORIGIN

Alignment Scores:

Pred. No.: 7.33 Length: 2557
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-20 (1-9) x AX961912 (1-2557)

Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9

Db 844 TTCCTTACCGGCAACACAGATGACCGTG 870

RESULT 8

AX961914 AX961914 2557 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 125 from Patent WO03104277.
DEFINITION AX961914
ACCESSION AX961914
VERSION AX961914.1 GI:40881324

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 125 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)

FEATURES

source 1..2557 Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06466.1"
/db_xref="GI:40881325"
/translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSSPSSSTS
PADFLASGAQPPPAERCPAAECSEAAATVKCVNRNLLEVPADLPYVNRNLFITGNQ
MTVLPAGAFARQPPPLADLEALNSGNHLKEVCAGAFEHLPGRLRLDLSHNPNTLSAF
VFAGSNVSAFSPLEELIINHIVPPDQQRQNGSFEGMVAFAALRSGLALRGL
TRLEASNHFLFPRDLAQLPSRLYLDLRNNSLSVLTYSFRNLTHLESLEHDNAL
KVLHNSTLAEMQGLAHVKVFLDNNPWCDQCYMADMAWLKETEVPDPKARLTCAFPK
MNRGLDLSNDDCDVLPQSLQTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"

CDS

1..2557
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
3108..3113
/clone_lib="Lambda Dash"
/bound_moiety="Sp1"
3114..3119
/bound_moiety="Sp1"
3124..5779
/gene="574"
3124..3151
/gene="574"
3152..3450
/gene="574"
3451..5779
/gene="574"
3779..5059
/gene="574"
/codon_start=1
/product="574 oncofetal trophoblast glycoprotein"
/protein_id="CAA09931.1"
/db_xref="GI:3805949"
/db_xref="GOA:Q9Z0L0"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProtKB/T-EMBL:Q9Z0L0"
/translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSSPSSSTS
PADFLASGAQPPPAERCPAAECSEAAATVKCVNRNLLEVPADLPYVNRNLFITGNQ
MTVLPAGAFARQPPPLADLEALNSGNHLKEVCAGAFEHLPGRLRLDLSHNPNTLSAF
VFAGSNVSAFSPLEELIINHIVPPDQQRQNGSFEGMVAFAALRSGLALRGL
TRLEASNHFLFPRDLAQLPSRLYLDLRNNSLSVLTYSFRNLTHLESLEHDNAL
KVLHNSTLAEMQGLAHVKVFLDNNPWCDQCYMADMAWLKETEVPDPKARLTCAFPK
MNRGLDLSNDDCDVLPQSLQTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"

ORIGIN

Alignment Scores:

Pred. No.: 7.33 Length: 2557
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-20 (1-9) x AX961914 (1-2557)

Qy 1 PheLeuThrGlyAsnGlnMetThrVal 9

Db 844 TTCCTTACCGGCAACACAGATGACCGTG 870

RESULT 9

MMU012160 MMU012160 7942 bp DNA linear ROD 15-APR-2005
LOCUS Mus musculus 574 oncofetal trophoblast glycoprotein gene.
DEFINITION

ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 574 gene; 574 oncofetal trophoblast glycoprotein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
TITLE Organisation of the mouse and human 574 oncofetal leucine-rich glycoprotein genes and expression in foetal and adult murine tissues

JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)

PUBMED

10366710

REFERENCE

2 (bases 1 to 7942)

AUTHORS

Myers, K.A.

TITLE

Direct Submission

JOURNAL

Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson Institute for Cancer Research, Christie Hospital, Wilmslow Road, Manchester, M20 9BX, UK

FEATURES

Location/Qualifiers

1..7942

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="129/Sv"

/db_xref="taxon:10090"

/clone_lib="Lambda Dash"

3108..3113

/bound_moiety="Sp1"

3114..3119

/bound_moiety="Sp1"

3124..5779

/gene="574"

3124..3151

/gene="574"

3152..3450

/gene="574"

3451..5779

/gene="574"

3779..5059

/gene="574"

/codon_start=1

/product="574 oncofetal trophoblast glycoprotein"

/protein_id="CAA09931.1"

/db_xref="GI:3805949"

/db_xref="GOA:Q9Z0L0"

/db_xref="InterPro:IPR000372"

/db_xref="InterPro:IPR000483"

/db_xref="InterPro:IPR001611"

/db_xref="InterPro:IPR003591"

/db_xref="MGI:1341264"

/db_xref="UniProtKB/T-EMBL:Q9Z0L0"

/translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSSPSSSTS
PADFLASGAQPPPAERCPAAECSEAAATVKCVNRNLLEVPADLPYVNRNLFITGNQ
MTVLPAGAFARQPPPLADLEALNSGNHLKEVCAGAFEHLPGRLRLDLSHNPNTLSAF
VFAGSNVSAFSPLEELIINHIVPPDQQRQNGSFEGMVAFAALRSGLALRGL
TRLEASNHFLFPRDLAQLPSRLYLDLRNNSLSVLTYSFRNLTHLESLEHDNAL
KVLHNSTLAEMQGLAHVKVFLDNNPWCDQCYMADMAWLKETEVPDPKARLTCAFPK
MNRGLDLSNDDCDVLPQSLQTSYVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"

3779..3865

/gene="574"

3866..5056

/gene="574"

/product="574 oncofetal trophoblast glycoprotein"

5713..5718

/gene="574"

5759..5764

/gene="574"

/sig_peptide

/mat_peptide

/polyA_signal

/polyA_signal

/ORIGIN

Alignment Scores:

Pred. No.: 23.5 Length: 7942
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-20 (1-9) x MM0012160 (1-7942)

QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
 |||||
 Db 4067 TTCCTTACCGGCACCATGACCGTG 4093

RESULT 10
 AC158516/c
 LOCUS AC158516 167046 bp DNA linear ROD 21-JUN-2005
 DEFINITION Mus musculus BAC clone RP24-511A23 from chromosome 9, complete sequence.

ACCESSION AC158516 AC117768
 VERSION AC158516.2 GI:63025421
 KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 167046)
 Adams, S., Cotton, M. and Haglund, K.

TITLE The sequence of Mus musculus BAC clone RP24-511A23

JOURNAL Unpublished (2001)

REFERENCE 2 (bases 1 to 167046)
 Wilson, R.K.

AUTHORS

TITLE Direct Submission

JOURNAL Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 167046)
 Wilson, R.K.

AUTHORS

TITLE Direct Submission

JOURNAL Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 167046)
 Wilson, R.K.

AUTHORS

TITLE Direct Submission

JOURNAL Submitted (21-JUN-2005) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On May 4, 2005 this sequence version replaced gi:61656412.

COMMENT

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu

Contact: submissions@wustl.edu

----- Summary Statistics

Center project name: M_BB0511A23

Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e. phred quality
 >30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone, fosmid clone or direct clone walk sequence.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to obtain the consensus sequence. The
 assembly was confirmed by restriction digest.

This finishing standard has slightly changed from the previous
 Human standard. Specifically, standards for regions of low sequence
 complexity (such as dinucleotide repeats and small unit tandem
 repeats) have been relaxed. These regions are very prevalent in the
 mouse genome, and the return on extended finishing efforts is
 minimal.

If a sequence meets the criteria of the above statement, it needs
 no comments or tags. If the criteria are not met, such as ambiguous
 bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The BAC library has been constructed by Pieter de Jong and
 coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
 and/or brain genomic DNA. The clone and detailed information can be
 obtained from Pieter de Jong and coworkers at http://www.chori.org

This sequence is the entire insert of the clone.

FEATURES

source

Location/Qualifiers
 1..167046
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="9"
 /clone="RP24-511A23"
 /clone_lib="RPCI-24"
 16685..16712

misc_feature

/note="Sequence derived from PCR product of genomic DNA"

unsure

31565..31779

/note="Unresolved simple sequence repeat."

unsure

46721..46808

/note="Unresolved simple sequence repeat."

unsure

142336..142347

/note="Sequence derived from one plasmid subclone."

ORIGIN

Alignment Scores:
 Pred. No.: 541 Length: 167046
 Score: 46.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-20 (1-9) x AC158516 (1-167046)

QY 1 PheLeuThrGlyAsnGlnMetThrVal 9

|||||
 Db 110550 TTCCTTACCGGCACCATGACCGTG 110524

RESULT 11

AC128294/c

LOCUS AC128294 210237 bp DNA linear HTG 19-NOV-2002

DEFINITION Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.

ACCESSION AC128294

VERSION AC128294.3 GI:25083347

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 210237)

AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abranzon, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anylebeche, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Cencer, A.,

Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, C.L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzshwa, L., Loulsegred, H., Lorado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mamoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathewney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwokeneme, O., Okwono, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, W., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Snead, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 210237)
Worley, K.C.

Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 210237)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265004.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZGV
Center clone name: CH230-176H20

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 201761 bases at least Q40
Consensus quality: 203921 bases at least Q30
Consensus quality: 205310 bases at least Q20
Estimated insert size: 205531; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 210237: contig of 210237 bp in length.

FEATURES
source
1 .210237
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-176H20"
1 .1142
/note="wgs_end_extension
clone_end:T7"
2177-144799
/note="clone_boundary
clone_end:T7
site:
end sequence: BH360464"
complement(206062..206961)
/note="clone_boundary
clone_end:Sp6
site:
end sequence: BH360465"
208907..210237
/note="wgs_end_extension
clone_end:Sp6"

misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Alignment Scores:
Pred. No.: 685 Length: 210237
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-20 (1-9) x ACL28294 (1-210237)
QY 1 PheLeuThrGlyAsnGlnMetThrVal 9
|||||
DB 111230 TTCTCTACCTGGCAACCAACGATGACCGTG 111204
|||||

RESULT 12
AC106962/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC106962
VERSION AC106962.5 GI:25139469
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 210237)
AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alabrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, J., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhray, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowig, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loulsegad, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Mlilosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoleneh, O., Okwunigbo, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Piazio, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 239076)
Worley, K.C.
Direct Submission
Submitted (14-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 239076)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GOPI
Center clone name: CH230-87110

Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 228642 bases at least Q40
Consensus quality: 23289 bases at least Q30
Consensus quality: 234041 bases at least Q20
Estimated insert size: 231522; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 234710: contig of 234710 bp in length
234711 234810: gap of unknown length
234811 235924: contig of 1114 bp in length
235925 236024: gap of unknown length
236025 237314: contig of 1290 bp in length
237315 237414: gap of unknown length
237415 239076: contig of 1662 bp in length.

FEATURES

Location/Qualifiers
1..239076
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-87110"
234711..234810
/estimated_length=unknown
235925..236024
/estimated_length=unknown
237315..237414
/estimated_length=unknown

ORIGIN

Alignment Scores:
Pred. No.: 782 Length: 239076
Score: 46.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-20 (1-9) x AC106962 (1-239076)

QY 1 PheLeuThrGlyAsnGlnMetThrVal 9

Db 16441 TTCTCACTGGCAACCAAGATGACCGTG 16415

RESULT 13

AC130132

LOCUS

DEFINITION

AC130132

AC130132.3 GI:30520865

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

COMMENT

AC130132 189229 bp DNA linear HTG 10-MAY-2003
Rattus norvegicus clone CH230-129E18, WORKING DRAFT SEQUENCE, 4
unordered pieces.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE AUTHORS

1 (bases 1 to 189229)
Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alebrooks,S., Amin,A., Anguiano,D., Anvalbechi,V., Aovagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Blawie,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,J., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.B., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Derramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Evans,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mathiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munida,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwankwelu,O., Okwuonu,G., Olarpunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puzoz,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,M., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villanana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,X., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

TITLE JOURNAL

2 (bases 1 to 189229)
Worley,K.C.
Direct Submission
Submitted (08-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE AUTHORS

3 (bases 1 to 189229)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GIAB

Center clone name: CH230-129E18

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 158286 bases at least Q40

Consensus quality: 161560 bases at least Q30

Consensus quality: 163846 bases at least Q20

Estimated insert size: 170094; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 184194: contig of 184194 bp in length

* 184195 184294: gap of unknown length

* 184295 185635: contig of 1341 bp in length

* 185636 185735: gap of unknown length

* 185736 186944: contig of 1209 bp in length

* 186945 187044: gap of unknown length

* 187045 189229: contig of 2185 bp in length.

FEATURES

source

1..189229

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-129E18"

46552..48982

/note="wgs_end_extension"

/clone_end:T7"

79767..103154

/note="clone_boundary"

/clone_end:T7"

site:ECORI

end_sequence:BH286393"

139780..141196

/note="wgs_contig"

complement(182186..183006)

/note="clone_boundary"

clone_end:Sp6

site:ECORI

end_sequence:BH286394"

184195..184294

/estimated_length=unknown

185636..185735

/estimated_length=unknown

186945..187044

/estimated_length=unknown

ORIGIN

Alignment Scores:

Pred. No.: 2.86e+03

Score: 43.00

Percent Similarity: 100.0%

Length: 189229

Matches: 8

Conservative: 1

```

Best Local Similarity: 88.9% Mismatches: 0
Query Match: 93.5% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-20 (1-9) x AC130132 (1-189229)

QY 1 PheLeuThrGlyAenGlnMetThrVal 9
|||||
Db 179438 TTCTTACAGGACCACTGACTATG 179464

RESULT 14
AC128563 228414 bp DNA linear HTG 19-SEP-2002
LOCUS Rattus norvegicus clone CH230-406A6, *** SEQUENCING IN PROGRESS
DEFINITION ****, 4 unordered pieces.
ACCESSION AC128563
VERSION AC128563.2 GI:23196177
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
1 (bases 1 to 228414)
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Daviila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Diviya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebraeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandes, B., Hines, S., Hladun, S.L., Hodgson, A., Hoggues, M.,
Hollins, R., Howells, S., Hui, J., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Louised, H., Lozano, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwokeleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Soralle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weis, R., Smith, D.R., Hoit, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 228414)
Worley, K.C.

Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
JOURNAL

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
sequence may extend beyond the ends of the clone and there may be
contigs that consist entirely of whole genome shotgun sequence
reads. Both end sequences and whole genome shotgun sequence only
contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYBP
Center clone name: CH230-406A6

----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 206214 bases at least Q40
Consensus quality: 209397 bases at least Q30
Consensus quality: 211141 bases at least Q20
Estimated insert size: 227271; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 206729: contig of 206729 bp in length
* 206730 206829: gap of unknown length
* 206830 224742: contig of 17913 bp in length
* 224743 224842: gap of unknown length
* 224843 226356: contig of 1514 bp in length
* 226357 226456: gap of unknown length
* 226457 228414: contig of 1958 bp in length.

FEATURES

source

1..228414
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-406A6"
1..1505
/note="wgs contig"

misc_feature

2420..4597
/note="wgs contig"

misc_feature

200269..202348
/note="wgs contig"

misc_feature

204485..206729
/note="wgs contig"

gap

206730..206829
/estimated_length=unknown

misc_feature

206830..208310
/note="wgs contig"

```

Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics -----
Center project name: M_BB0351117
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC121871.

Location/Qualifiers
1..179668
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="10"
/map="10"
/clone="RP24-351117"
/clone_lib="RPCI-24"
1..46
/rpt_family="ERVK"
_region 49..271
/rpt_family="MER2_type"
_region 360..462
/rpt_family="MER1_type"
_region 1825..2101
/rpt_family="ERVL"
_region 2145..2288
/rpt_family="Alu"
_region 2299..2444
/rpt_family="ERVL"
_region 3239..3384
/rpt_family="ERVL"
_region 3385..3779
/rpt_family="MaLR"
_region 3780..3838
/rpt_family="ERVL"
_region 4249..4475
/rpt_family="MER1_type"
_region 4755..5164
/rpt_family="MaLR"
_region 5323..5682
/rpt_family="MaLR"
_region 5869..6060
/rpt_family="ERVL"
_region 6476..6874
/rpt_family="L1"
_region 7871..8002

```

repeat_region /rpt_family="Alu" 9253. .9627
repeat_region /rpt_family="L1" 9830. .9956
repeat_region /rpt_family="Alu" 10866. .10363
repeat_region /rpt_family="MaLR" 10376. .10474
repeat_region /rpt_family="MaLR" 10479. .10614
repeat_region /rpt_family="Alu" 10690. .10901
repeat_region /rpt_family="MaLR" 11022. .11289
repeat_region /rpt_family="B4" 11534. .11675
repeat_region /rpt_family="Alu" 13608. .13776
repeat_region /rpt_family="B2" 14130. .14273
repeat_region /rpt_family="MIR" 15625. .15697
repeat_region /rpt_family="Alu" 15715. .16237
repeat_region /rpt_family="L1" 16238. .16382
repeat_region /rpt_family="B2" 16448. .16544
repeat_region /rpt_family="B4" 17860. .18123
repeat_region /rpt_family="MaLR" 18333. .18539
repeat_region /rpt_family="B4" 18855. .19224
repeat_region /rpt_family="MaLR" 19229. .19328
repeat_region /rpt_family="MaLR" 19814. .20015
repeat_region /rpt_family="B2" 20281. .20620
repeat_region /rpt_family="MaLR" 20621. .20945
repeat_region /rpt_family="MaLR" 20946. .20981
repeat_region /rpt_family="MaLR" 21348. .21534
repeat_region /rpt_family="B2" 21744. .22005
repeat_region /rpt_family="L1" 22055. .22181
repeat_region /rpt_family="Alu" 23118. .23185
repeat_region /rpt_family="Alu" 23186. .23209
repeat_region /rpt_family="B4" 23704. .23831
repeat_region /rpt_family="Alu" 24302. .24442
repeat_region /rpt_family="L2" 24935. .25009
repeat_region /rpt_family="ERV1" 25102. .25201
repeat_region /rpt_family="ERV1" 25399. .25486
repeat_region /rpt_family="MIR" 25652. .25760
repeat_region /rpt_family="Alu" 25856. .25935
repeat_region /rpt_family="MER2_type" 25939. .26042
repeat_region /rpt_family="MER2_type" 26275. .26381
repeat_region /rpt_family="B4"

repeat_region 26320. .26391 /rpt_family="Alu"
repeat_region 26392. .26532 /rpt_family="B4"
repeat_region 26776. .26984 /rpt_family="ERV1"
repeat_region 27451. .27631 /rpt_family="B2"
repeat_region 28016. .28177 /rpt_family="L1"
repeat_region 28186. .28519

Alignment Scores:
Pred. No.: 4.52e+03 Length: 179668
Score: 42.00 Matches: 8
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.3% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-20 (1-9) x AC124399 (1-179668)

QY 1 PheLeuThrGlyAsnGlnMetThr 8
DB 140708 TTCTTGACAGGAATCAGATGACA 140685

Search completed: May 27, 2006, 19:35:20
Job time : 3381.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-19
Perfect score: 44
Sequence: 1 NLEVPADL 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB spo0/US10774176/runat 26052006 091441 24976/app query.fasta 1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT5-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US10774176 @CGN 1 1 2389 @runat 26052006 091441 24976 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 8.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*
- 15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	44	100.0	1281	3	AAA27059	Mouse 5T4
2	44	100.0	2557	12	Adi26160	Human cdn
3	44	100.0	2557	12	Adi26158	Human cdn

C	4	44	100.0	2557	12	ADO35939	Ado35939 Novel mou
	5	39	88.6	343	8	ABX40605	Abx40605 Bovine ES
	6	39	88.6	1260	6	ABK87175	Abk87175 cDNA enco
	7	39	88.6	1260	10	ADB97513	Adb97513 Feline 5T
	8	39	88.6	1260	10	ADB97452	Adb97452 DNA enco
	9	39	88.6	1263	4	AAF89736	Aaf89736 Nucleotid
	10	39	88.6	1263	6	ABK87174	Abk87174 cDNA enco
	11	38	86.4	295096	11	ACN44068	Acn44068 Mouse gen
	12	37	84.1	1641	13	ACN39471	Acn39471 Tumour-as
	13	37	84.1	1753	14	AED22740	Aed22740 Human TIM
	14	37	84.1	3627	14	ADV43825	Adv43825 Human psy
	15	37	84.1	3627	14	AED22737	Aed22737 Human TIM
	16	37	84.1	4366	14	AED22736	Aed22736 Human TIM
	17	37	84.1	4414	3	AAA14684	Aaa14684 DNA enco
	18	37	84.1	4414	3	AAA14683	Aaa14683 DNA enco
	19	37	84.1	4414	3	AAA14685	Aaa14685 DNA enco
	20	37	84.1	4414	3	AAA14682	Aaa14682 DNA enco
	21	37	84.1	4414	3	AAA14687	Aaa14687 Variant D
	22	37	84.1	4417	3	AAZ34632	Aaz34632 Human cdn
	23	37	84.1	4417	3	AAA14678	Aaa14678 DNA enco
	24	37	84.1	4417	3	AAA14686	Aaa14686 Variant D
	25	37	84.1	4417	3	AAA14679	Aaa14679 DNA enco
	26	37	84.1	4417	3	AAA14680	Aaa14680 DNA enco
	27	37	84.1	4417	3	AAA14681	Aaa14681 DNA enco
	28	37	84.1	4417	14	AED22739	Aed22739 Human TIM
	29	37	84.1	4417	14	AED22738	Aed22738 Human TIM
	30	37	84.1	4427	13	ADR83429	Adr83429 Human tim
	31	37	84.1	4427	14	AED22735	Aed22735 Human TIM
	32	36	81.8	1143	13	ADS56356	Ad56356 Bacterial
C	33	36	81.8	2715	4	ABL23835	Ab123835 Drosophil
	34	36	81.8	3768	4	ABL05440	Ab105440 Drosophil
	35	36	81.8	5893	4	ABL23834	Ab123834 Drosophil
	36	36	81.8	110000	6	ABX08336_02	Continuation (3 of
C	37	36	81.8	110000	10	ACF67367_17	Continuation (18 o
	38	36	81.8	110000	12	ADJ25985_02	Continuation (3 of
	39	36	81.8	110000	12	ADN97989_02	Continuation (3 of
	40	36	81.8	110000	12	ADOS0281_02	Continuation (3 of
	41	36	81.8	110000	14	AE85185_02	Continuation (3 of
C	42	36	81.8	210710	10	ACF65380	Acf65380 Photornab
C	43	36	81.8	347001	12	ADP43517	Adp43517 Human MAD
	44	35	79.5	505	14	ACL56146	Ac156146 Human col
C	45	35	79.5	619	11	ADJ11918	Adj11918 Wheat cdn

ALIGNMENTS

RESULT 1
AAA27059
ID AAA27059 standard; DNA; 1281 BP.
XX
AC AAA27059;
XX
DT 22-AUG-2000 (first entry)
XX
DE Mouse 5T4 tumour-associated antigen gene.
XX
KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX Mus musculus.
XX
XX WO200029428-A2.
XX
XX
PD 25-MAY-2000.
XX
XX 18-NOV-1999; 99WO-GB003859.
XX
XX 18-NOV-1998; 98GB-00025303.
PR 27-JAN-1999; 99GB-00001739.
PR 30-JUL-1999; 99GB-00017995.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll MW, Myers KA;
 XX WPI; 2000-387735/33.
 XX
 PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.
 XX
 PS Example 2; Page 78; 79pp; English.
 XX
 CC The present sequence encodes the mouse 5T4 tumour-associated antigen
 CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
 CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The 5T4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject. The
 CC present sequence appears in GenBank at accession number AJ012160
 XX
 SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 17.9 Length: 1281
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-19 (1-9) x AAA27059 (1-1281)
 QY 1 AsnLeuLeuGluValProAlaAspLeu 9
 Db 241 AACCTGCTGGAGGTGCCGCGGATCTA 267
 RESULT 2
 ADI26160
 ID ADI26160 standard; cDNA; 2557 BP.
 XX
 AC ADI26160;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #63.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 XX WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 PF 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00137505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 XX (ASAH) ASAH KASEI KK.
 XX

PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX WPI; 2004-122214/12.
 DR P-PSDB; ADI26161.
 XX
 PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 PS Claim 4; SEQ ID NO 125; 1368pp; English.
 XX
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 39.1 Length: 2557
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-19 (1-9) x ADI26160 (1-2557)
 QY 1 AsnLeuLeuGluValProAlaAspLeu 9
 Db 796 AACCTGCTGGAGGTGCCGCGGATCTA 822
 RESULT 3
 ADI26158
 ID ADI26158 standard; cDNA; 2557 BP.
 XX
 AC ADI26158;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #62.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 XX WO2003104277-A2.
 XX
 PD 18-DEC-2003.

XX 05-JUN-2003; 2003WO-JP007123.
PF 05-JUN-2002; 2002JP-00164257.
XX 06-JUN-2002; 2002US-0385912P.
PR 26-DEC-2002; 2002JP-00377326.
PR 27-DEC-2002; 2002US-0436467P.
PR 15-MAY-2003; 2003JP-00137505.
PR 16-MAY-2003; 2003US-0470836P.
XX
PA (ASAH) ASahi KASEI KK.
XX
XX Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
PI WPI; 2004-122214/12.
XX P-PSDB; ADI26159.
XX
XX New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX
XX Claim 4; SEQ ID NO 123; 1368pp; English.
PS
XX
XX The invention relates to a purified protein promoting signal transducer
CC and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infections disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.
XX
SQ Sequence 2557 BP; 512 A; 730 C; 696 G; 619 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 39.1 Length: 2557
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-19 (1-9) x ADI26158 (1-2557)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9
Db 796 AACCTGCTGGAGGTGCCGGCGGATCTA 822

RESULT 4
ID ADO35939/c standard; DNA; 2557 BP.
XX
AC ADO35939;
XX
XX 26-AUG-2004 (first entry)
XX
XX Novel mouse gene sequence #612.
DE
XX

KW mouse; murine; cancer; psoriasis; ulcerative colitis; inflammation;
KW ischaemic heart disease; thrombosis; immune disorder; bacterial disorder;
KW viral disorder; ds; gene.
XX
OS Mus sp.
XX
PN WO2004046310-A2.
XX
PD 03-JUN-2004.
XX
PF 24-OCT-2003; 2003WO-US033948.
XX
PR 15-NOV-2002; 2002US-0426916P.
PR 04-DEC-2002; 2002US-0431158P.
PR 05-DEC-2002; 2002US-0431445P.
PR 05-DEC-2002; 2002US-0431606P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485359P.
PR 08-AUG-2003; 2003US-0493332P.
PR 08-AUG-2003; 2003US-0493356P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Hayashizaki Y, Kamiya M;
PI WPI; 2004-431966/40.
XX
XX New mouse nucleic acid molecules and polypeptides, useful for treating
PT cancer, psoriasis, ulcerative colitis, inflammation, ischemic heart
PT disease or thrombosis.
XX
XX Claim 1; SEQ ID NO 612; 263pp; English.
PS
XX The invention comprises 744 novel mouse DNA sequences (genes). The DNA
CC sequences of the invention are useful for treating cancer, psoriasis,
CC ulcerative colitis, inflammation, ischaemic heart disease, thrombosis,
CC immune disorders, bacterial disorders and viral disorders. The present
CC nucleic acid represents a mouse DNA sequence of the invention. NOTE: The
CC present DNA sequence is not shown in the specification, but has been
CC retrieved from the WIPO website.
XX
SQ Sequence 2557 BP; 610 A; 794 C; 688 G; 465 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 39.1 Length: 2557
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-774-176-19 (1-9) x ADO35939 (1-2557)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9
Db 105 AACCTGCTGGAGGTGCCGGCGGATCTA 79

RESULT 5
ABX40605
ID ABX40605 standard; cDNA; 343 BP.
XX
XX ABX40605;
XX
XX 20-FEB-2003 (first entry)
DT
XX
XX Bovine EST associated with lactation/muscle/fat deposition #5770.
DE
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
XX

OS Bos Taurus.
 XX US2002137139-A1.
 XX
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-00960352.
 XX
 XX 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 XX (BYAT// BYATT J C.
 PA (MATH// MATHIALAGAN N.
 PA (TAON// TAO N.
 PA (WARR// WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 5770; 245pp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non- translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docid=20020137139
 XX

XX Sequence 343 BP; 40 A; 146 C; 108 G; 49 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 47.8 Length: 343
 Score: 39.00 Matches: 8
 Percent Similarity: 88.9% Conservative: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 88.6% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-19 (1-9) x ABX40605 (1-343)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9
 Db 131 AACTTGACCGAGGTGCCGCGACCTG 157

RESULT 6
 ABK87175
 ID ABK87175 standard; cDNA; 1260 BP.
 XX

AC ABK87175;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE cDNA encoding feline oncofoetal leucine-rich glycoprotein, 574.
 XX
 KW Feline; cat; oncofoetal leucine-rich glycoprotein; 574; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX
 OS Felis sp.
 XX
 XX Key Location/Qualifiers
 FT 1..1260
 FT /*tag= a
 FT /product= "574 protein"
 XX
 PN WO200238612-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 13-NOV-2001; 2001WO-GB005004.
 XX
 PR 13-NOV-2000; 2000WO-GB004317.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 PA
 PI Myers K, Drury N, Carroll M;
 XX WPI; 2002-557449/59.
 DR P-PSDB; AAU98694.
 XX
 PT Novel canine or feline 574 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.
 XX
 PS Claim 4; Page 68; 68pp; English.
 XX
 CC The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 574, and the
 CC polynucleotide sequences encoding them. The 574 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 574 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 574
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes feline 574 protein

XX Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 207 Length: 1260
 Score: 39.00 Matches: 8
 Percent Similarity: 88.9% Conservative: 0
 Best Local Similarity: 88.9% Mismatches: 1
 Query Match: 88.6% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-19 (1-9) x ABK87175 (1-1260)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9
 Db 238 AACCTGACCGAGGTGCCGCGACCTG 264

```

RESULT 7
ADB97513
ID ADB97513 standard; DNA; 1260 BP.
XX
XX AC ADB97513;
XX
XX DT 04-DEC-2003 (first entry)
XX
XX DE Feline 5T4 antigen DNA.
XX
XX Major Histocompatibility Complex class I peptide epitope; MHC;
KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
KW cytostatic; cancer; feline; gene; ds.
XX
XX OS Unidentified.
XX
XX Key Location/Qualifiers
FH 1..1260
FT CDS
FT /*tag= a
FT /product= "Feline 5T4 antigen protein"
XX
XX WO2003068816-A1.
XX
XX 21-AUG-2003.
XX
XX 13-FEB-2003; 2003WO-GB000670.
XX
XX 13-FEB-2002; 2002GB-00003419.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll M, Kingman S, Redchenko I;
XX
XX WPI; 2003-637141/60.
XX
XX P-PSDB; ADB97520.
XX
XX New major histocompatibility complex class I peptide epitopes from human
PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
PT disease, particularly cancer.
XX
XX Disclosure; Page 67; 73pp; English.
XX
XX The invention relates to a novel Major Histocompatibility Complex (MHC)
CC class I peptide epitope of the 5T4 antigen. The invention further
CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
CC comprising the above; a method for treating and/or preventing a disease
CC in a subject by administering the vaccine; an agent capable of binding
CC specifically to the 5T4 epitope and/its encoding nucleic acid; a method
CC comprising detecting the presence of the 5T4 epitope or its encoding
CC nucleic acid in a subject; and a T cell line or clone capable of
CC specifically recognising the 5T4 epitope in conjunction with an MHC class
CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
CC system or cell is useful in the prevention and/or treatment of a disease,
CC particularly cancer. The detection method is useful for diagnosing or
CC monitoring the progression of a cancerous disease, and for detecting the
CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
CC is useful in the manufacture of a medicament for treating and/or
CC preventing a disease. This polynucleotide sequence represents the feline
CC 5T4 antigen coding DNA of the invention.
XX
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 207 Length: 1260
Score: 39.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 88.6% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-19 (1-9) x ADB97513 (1-1260)
QY 1 AsnLeuLeuGluValProAlaAspLeu 9
Db 238 AACCTGACCGAGGTGCCCGGACCTG 264

RESULT 8
ADB97452
ID ADB97452 standard; DNA; 1260 BP.
XX
XX AC ADB97452;
XX
XX DT 04-DEC-2003 (first entry)
XX
XX DE DNA encoding feline 5T4 protein.
XX
XX gene; ds; feline; Major Histocompatibility Complex class II; MHC;
KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
XX
XX OS Unidentified.
XX
XX Key Location/Qualifiers
FH 1..1260
FT CDS
FT /*tag= a
FT /product= "Feline 5T4 antigen protein"
XX
XX WO2003068815-A2.
XX
XX 21-AUG-2003.
XX
XX 13-FEB-2003; 2003WO-GB000618.
XX
XX 13-FEB-2002; 2002GB-00003420.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Carroll M, Harrop R, Kingman S;
XX
XX WPI; 2003-663795/62.
XX
XX P-PSDB; ADB97455.
XX
XX New Major Histocompatibility Complex class II peptide epitope of 5T4,
PT useful for manufacturing a medicament for diagnosing, preventing and/or
PT treating a disease, e.g. cancer.
XX
XX Disclosure; Page 49; 63pp; English.
XX
XX The invention relates to a Major Histocompatibility Complex (MHC) class
CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
CC clone has a cytostatic activity, as it is useful in manufacturing a
CC medicament for preventing and/or treating a disease, particularly cancer.
CC The methods are useful for detecting T-cells capable of specifically
CC recognising a peptide epitope in conjunction with an MHC molecule, for
CC diagnosing or monitoring the progression of a cancerous disease, or for
CC detecting the presence of a peptide or nucleic acid using an agent. The
CC MHC class II peptide epitope of the invention can be used in gene therapy
CC or as part of a vaccine. This polynucleotide sequence represents the DNA
CC coding for the feline 5T4 protein.
XX
XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 207 Length: 1260
Score: 39.00 Matches: 8
Percent Similarity: 88.9% Conservative: 0
Best Local Similarity: 88.9% Mismatches: 1
Query Match: 88.6% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-19 (1-9) x ADB97452 (1-1260)

```

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
DB 238 AACCTGACCGAGGTGCCCGGACCTG 264

RESULT 9
ID AAF89736
AC AAF89736 standard; DNA; 1263 BP.
XX
XX
XX 23-JUL-2001 (first entry)
DE Nucleotide sequence of canine 5T4 protein.
XX
XX Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
KW hypersensitivity; autoimmune disease; central nervous system disorder;
KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
KW cardiovascular disease; gastrointestinal disorder; infection; diabetes;
KW Helicobacter-related disease; immune disorder; ss.
XX
XX Canis sp.
XX
XX
XX Key Location/Qualifiers
CDS 1..1263
FT /*tag= a
FT /product= "5T4"
XX
XX WO200136486-A2.
XX
XX 25-MAY-2001.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX 18-NOV-1999; 99WO-GB003859.
XX 15-FEB-2000; 2000GB-00003527.
XX 02-MAR-2000; 2000GB-00005071.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
PI Myers KA;
XX
XX WPI; 2001-343805/36.
DR P-PSDB; AAB83839.
XX
XX Use of single chain antibody capable of recognizing a disease associated
PT molecule for manufacturing a medicament for preventing and/or treating a
PT disease condition associated with disease associated molecule.
XX
XX Disclosure; Fig 26; 118pp; English.

CC The specification describes the use of a single chain antibody (ScFv),
CC which is capable of recognizing a disease associated molecule in the
CC manufacture of a medicament for the prevention and treatment of a disease
CC condition. The ScFv antibody is useful in the manufacture of a
CC medicament, for affecting a disease in vivo, for preparing a
CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
CC treatment of a disease. The ScFv antibody is also useful for treating
CC inflammatory diseases including arthritis, hypersensitivity, autoimmune
CC diseases, cancers, central nervous system disorders including Parkinson's
CC disease, periodontal disease, cardiopulmonary diseases, cardiovascular
CC diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-
CC related diseases, and other immune disorders. The present sequence
CC encodes a 5T4 protein, which is used to produce ScFv of the invention
XX
SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores: 208 Length: 1263
Pred. No.: 39.00 Matches: 8
Score: 88.9% Conservative: 0
Percent Similarity: 88.9% Mismatches: 1
Best Local Similarity: 88.9% Indels: 0
Query Match: 88.6%

DB: 4 Gaps: 0
US-10-774-176-19 (1-9) x AAF89736 (1-1263)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
DB 241 AACCTGACCGAGGTGCCCGGACCTG 267

RESULT 10
ID ABK87174 standard; cDNA; 1263 BP.
XX
XX AC ABK87174;
XX
XX 07-OCT-2002 (first entry)
DE cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
XX
XX Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX
XX Canis sp.
XX
XX Key Location/Qualifiers
CDS 1..1263
FT /*tag= a
FT /product= "5T4 protein"
XX
XX WO200238612-A2.
XX
XX 16-MAY-2002.
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Myers K, Drury N, Carroll M;
PI WPI; 2002-557449/59.
DR P-PSDB; AAU98693.
XX
XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 1; Page 67; 68pp; English.

CC The present invention relates to the isolation of canine and feline
CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 5T4 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 5T4
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus(es).
CC The present sequence encodes canine 5T4 protein

SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores: 208 Length: 1263
Pred. No.: 39.00 Matches: 8
Score: 88.9% Conservative: 0
Percent Similarity: 88.9% Mismatches: 1
Best Local Similarity: 88.9% Indels: 0
Query Match: 88.6%

DB: 13 Gaps: 0

US-10-774-176-19 (1-9) x ACN39471 (1-1641)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
||||:|||||

Db 556 AATATTCTCCATGTCCTCCAGCTGACCTT 582

RESULT 13
AED22740
ID AED22740 standard; cDNA; 1753 BP.

XX
AC AED22740;

XX
DT 15-DEC-2005 (first entry)

XX
DE Human TIMELESS homolog, cDNA SEQ ID 632.

XX
XX RNA interference; gene silencing; ss; carcinoma; cytostatic; apoptosis;

KW chromosome stabilization; cancer; DNA damage.

XX
OS Homo sapiens.

XX
PN WO2005097189-A1.

XX
PD 20-OCT-2005.

XX
PF 08-APR-2005; 2005WO-JP006914.

XX
PR 09-APR-2004; 2004JP-00115404.

XX
PA (GENE-) GENE CARE RES INST CO LTD.

XX
PI Takagi M, Futami K, Shimamoto A, Furuichi Y;

XX
XX WPI; 2005-725737/74.

XX
PT Cancer cell-specific apoptosis inducing agent useful as anticancer agent,

PT comprises as active ingredient, compound that inhibits stabilization of

PT chromosome or compound that suppresses expression of gene e.g. APE2,

PT BRCA1, Cdc7, NEIL2.

XX
PS Claim 5; SEQ ID NO 632; 181bp; Japanese.

XX
CC The invention relates to a cancer cell-specific apoptosis inducing agent

CC comprising, as an active ingredient, a compound which inhibits

CC stabilization of a chromosome, or a compound which suppresses the

CC expression of a gene chosen from APE2, ATR, BRCA1, Chk1, Cdc5, Cdc6,

CC Cdc7, Cdc45, Cdt1, CSA, CSB, Ctf18, DDB1, DDB2, DNA2, DUT, Elg1, Endov,

CC Epl1, Exonuclease1, FBH1, FEN1, Geminin, Hus1, KNTC2 (NDC80), Ku80,

CC Ligase1, Mad2, MBD4, Mcm3, Mcm4, Mcm5, Mcm7, Mcm8, Mcm10, MGMT,

CC MLH3, Mre11, Mre11A, Mre11B, NBS1, NBS1L1, NBS1L2, NBS1L3, NBS1L4, NBS1L5,

CC PARP1, PCNA, Pif1, PMS1, PMS2, PNLK, Polap180, Pola p70, Pola Sppl

CC (Prim2a), Polb, Pold p125, Pole Dpb3, Pole Dpb4, Pole Pol2, Poli, Poll,

CC Polm, Psf1, Psf2, Psf3, Rad1, Rad18, Rad23A, Rad23B, Rad51, Rad51D,

CC Rad54, Rad6A, RPA34, RPA70, Scc1, Scc3, Sir2, SirT1 (Sirtuin), TD, TDPI,

CC TIMELESS, Tin2, Topoisomerase I, Topoisomerase IIIa, Topoisomerase IIb,

CC Ubc13, UNG, XAB2, XPC, XPF, XPG, Xrcc2 and Xrcc4. Also included are an

CC anticancer agent comprising (I) as an active ingredient, screening (M1)

CC cancer cell specific apoptosis inducing agents, and manufacturing (I) as

CC a pharmaceutical composition (involving screening the compound by (M1)

CC and mixing the compound with a carrier). In (I), the compound which

CC inhibits the stabilization of a chromosome, is an agent capable of

CC inhibiting a human chromosome instability disease related gene,

CC chromosome replication of DNA, chromosome DNA reproduction reaction, DNA

CC damage check point, sister chromatid aggregation or isolation, base

CC excision repair, mismatch excision repair, nucleotide excision repair,

CC homologous-recombination restoration, non-homologous terminal binding

CC restoration, double-stranded DNA-cleavage repair, DNA post replication

CC repair (DNA damage tolerance), DNA crosslinking damage repair, DNA-

CC protein crosslinking damage repair, DNA polymerase, nuclease, nucleotide

CC purification, chromatin structure maintenance and telomere structure

CC maintenance. The compound, which suppresses the expression of the genes

CC is a double-stranded RNA (siRNA, short interfering RNA), antisense

CC oligonucleotide, a variant protein having dominant negative character

CC with respect to protein encoded by the gene, an antibody and a low

CC molecular weight compound which binds to the protein encoded by the gene.

CC The compounds are useful as anticancer agents. (M1) is useful for

CC manufacturing a pharmaceutical composition comprising the compound. The

CC compound is useful for elucidating the mechanism of inducing cancer cell

CC specific apoptosis and is highly cancer cell specific, has few or no side

CC effects as an anticancer agent and damage produced in chromosomal DNA is

CC repaired. The present sequence is a cDNA representing a gene target for

CC the compounds of the invention.

SQ Sequence 1753 BP; 407 A; 457 C; 480 G; 409 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	808	Length:	1753
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	84.1%	Indels:	0
DB:	14	Gaps:	0

US-10-774-176-19 (1-9) x AED22740 (1-1753)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
||||:|||||

Db 648 AATATTCTCCATGTCCTCCAGCTGACCTT 674

RESULT 14

ADV43825

ID ADV43825 standard; cDNA; 3627 BP.

XX
AC ADV43825;

XX
DT 10-MAR-2005 (first entry)

DE Human psychoneuroendocrine-immune expressed sequence tag SEQ ID NO 1453.

KW microarray; psychoneuroendocrine-immune; chronic fatigue;

KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;

KW cancer; neoplasm; infection; expressed sequence tag; ss.

XX
OS Homo sapiens.

XX
PN WO2004108899-A2.

XX
PD 16-DEC-2004.

XX
PF 04-JUN-2004; 2004WO-US017686.

XX
PR 04-JUN-2003; 2003US-0475915P.

XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX
PI Nicholson A, Vernon SD;

XX
DR WPI; 2005-031682/03.

XX
PT New microarray comprising probes for genes involved in

PT psychoneuroendocrine-immune (PNI) activity, useful in diagnosing a

PT condition associated with PNI activity, e.g., inflammatory or infectious

PT diseases.

XX
PS Claim 1; SEQ ID NO 1453; 254pp; English.

XX
CC The invention relates to a new microarray which comprises probes for

CC genes involved in psychoneuroendocrine-immune (PNI) activity. The

CC microarray is useful in diagnosing a condition associated with PNI

CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,

CC cancer and infection. The present sequence represents a

CC psychoneuroendocrine-immune gene expressed sequence tag. Note the

CC specification mentions SEQ ID NO of up to 3314 but only sequences up to

CC SEQ ID NO 1829 are provided.

XX SQ Sequence 3627 BP; 926 A; 894 C; 1039 G; 768 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.83e+03	Length:	3627
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	84.1%	Indels:	0
DB:	14	Gaps:	0

US-10-774-176-19 (1-9) x ADV43825 (1-3627)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9
|||::||| |||||

Db 493 AATATTCTCCATGTCCCGACGCTT 519

RESULT 15

AED22737

ID AED22737 standard; cDNA; 3627 BP.

XX AC AED22737;

XX DT 15-DEC-2005 (first entry)

XX DE Human TIMELESS homolog, cDNA SEQ ID 629.

XX KW RNA interference; gene silencing; ss; carcinoma; cytostatic; apoptosis; chromosome stabilization; cancer; DNA damage.

XX KM

XX OS Homo sapiens.

XX PN WO2005097189-A1.

XX PD 20-OCT-2005.

XX PF 08-APR-2005; 2005WO-JP006914.

XX PR 09-APR-2004; 2004JP-00115404.

XX PA (GENE-) GENE CARE RES INST CO LTD.

XX PI Takagi M, Futami K, Shimamoto A, Furuichi Y;

XX DR WPI; 2005-725737/74.

XX PT Cancer cell-specific apoptosis inducing agent useful as anticancer agent, comprises as active ingredient, compound that inhibits stabilization of chromosome or compound that suppresses expression of gene e.g. APE2, BRCA1, Cdc7, NEIL2.

XX PS Claim 5; SEQ ID NO 629; 181pp; Japanese.

XX CC The invention relates to a cancer cell-specific apoptosis inducing agent comprising, as an active ingredient, a compound which inhibits stabilization of a chromosome, or a compound which suppresses the expression of a gene chosen from APE2, ATR, BRCA1, Chk1, Cdc5, Cdc6, Cdc7, Cdc45, Cdt1, CSA, CSB, Ctf18, DDB1, DDB2, DNA2, DUT, Elg1, Endov, Esp1, Exonuclease1, FBH1, FEN1, Geminin, Hse1, KNTC2 (NDC80), Ku80, Ligase1, Mad2, MBD4, Mcm3, Mcm4, Mcm5, Mcm6, Mcm7, Mcm8, Mcm10, MGMT, MLH3, Mms4, MPG, MSH2, Msh81, NBS1, NEIL2, NEIL3, NTH1, Orc1, Orc3, PARP1, PCNA, Pif1, PMS1, PMS2, PNK, Polap180, Pola p70, Pola Spp1 (Prim2a), Polb, Pold p125, Pole Dpb3, Pole Dpb4, Pole Pol2, Poli, Poll, Polm, Ppf1, Ppf2, Ppf3, Rad1, Rad18, Rad23A, Rad23B, Rad51, Rad51D, Rad54, Rad6A, RPA34, RPA70, Scc1, Scc3, Sir2, SIRT1 (Sirtuin), TD, TDP1, TIMELESS, Tin2, Topoisomerase I, Topoisomerase IIIa, Topoisomerase IIIB, Ubc13, UNG, XAB2, XPC, XPF, XPG, Xrcc2 and XRCC4. Also included are an anticancer agent comprising (I) as an active ingredient, screening (M1) cancer cell specific apoptosis inducing agents, and manufacturing (I) as a pharmaceutical composition (involving screening the compound by (M1) and mixing the compound with a carrier). In (I), the compound which inhibits the stabilization of a chromosome, is an agent capable of inhibiting a human chromosome instability disease related gene,

CC chromosome replication of DNA, chromosome DNA reproduction reaction, DNA damage check point, sister chromatid aggregation or isolation, base excision repair, mismatch excision repair, nucleotide excision repair, homologous-recombination restoration, non-homologous terminal binding, restoration, double-stranded DNA-cleavage repair, DNA post replication repair (DNA damage tolerance), DNA crosslinking damage repair, DNA-protein crosslinking damage repair, DNA polymerase, nuclease, nucleotide purification, chromatin structure maintenance and telomere structure maintenance. The compound, which suppresses the expression of the genes is a double-stranded RNA (siRNA, short interfering RNA), antisense oligonucleotide, a variant protein having dominant negative character with respect to protein encoded by the gene, an antibody and a low molecular weight compound which binds to the protein encoded by the gene. The compounds are useful as anticancer agents. (M1) is useful for manufacturing a pharmaceutical composition comprising the compound. The compound is useful for elucidating the mechanism of inducing cancer cell specific apoptosis and is highly cancer cell specific, has few or no side effects as an anticancer agent and damage produced in chromosomal DNA is repaired. The present sequence is a cDNA representing a gene target for the compounds of the invention.

XX SQ Sequence 3627 BP; 925 A; 893 C; 1039 G; 768 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	1.83e+03	Length:	3627
Score:	37.00	Matches:	7
Percent Similarity:	88.9%	Conservative:	1
Best Local Similarity:	77.8%	Mismatches:	1
Query Match:	84.1%	Indels:	0
DB:	14	Gaps:	0

US-10-774-176-19 (1-9) x AED22737 (1-3627)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9
|||::||| |||||

Db 493 AATATTCTCCATGTCCCGACGCTT 519

Search completed: May 27, 2006, 10:38:47
Job time : 388.5 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds
(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-19
Perfect score: 44
Sequence: 1 NLEVPADL 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/abss/ABSSWEB/spool/US10774176/runat_26052006_091443_24987/app_query.fasta_1
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss04
-USER=US10774176 @CCN_1_1_7524 @runat_26052006_091443_24987 -NCFU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sts.*
8: gb_sy.*
9: gb_un.*
10: gb_vl.*
11: gb_ov.*
12: gb_htg.*
13: gb_in.*
14: gb_on.*
15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	1281	2	BD249732 Polypepti
2	44	100.0	1281	2	AX025012 Sequence
3	44	100.0	1281	2	AX316087 Sequence

4	44	100.0	2333	6	AF063939	Rattus no
5	44	100.0	2361	6	BC087011	Rattus no
6	44	100.0	2423	6	BC058198	Mus muscu
7	44	100.0	2557	2	AX961912	Sequence
8	44	100.0	2557	2	AX961914	Sequence
9	44	100.0	7942	6	MMU012160	Mus muscu
10	44	100.0	167046	6	AC158516	Mus muscu
11	44	100.0	210237	12	AC128294	Rattus no
12	44	100.0	239076	12	AC106962	Rattus no
13	41	93.2	1125	13	AB003707	Maldane c
14	41	93.2	236951	12	AC098662	Rattus no
15	41	93.2	244451	12	AC125765	Rattus no
16	40	90.9	195185	6	AC124487	Mus muscu
17	39	88.6	586	7	BV291648	S23P26207
18	39	88.6	1260	2	AX467373	Sequence
19	39	88.6	1260	2	AX821533	Sequence
20	39	88.6	1260	2	AX821548	Sequence
21	39	88.6	1263	2	AX149553	Sequence
22	39	88.6	1263	2	AX467371	Sequence
23	39	88.6	77837	12	AC156143	Bos tauru
24	39	88.6	105936	13	AC084440	Caenorhab
25	39	88.6	110000	4	CR382128_21	Continuation (22 o
26	39	88.6	110000	15	CP000025_10	Continuation (11 o
27	39	88.6	110000	15	BA000002_06	Continuation (7 of
28	39	88.6	184731	12	AC165022	Bos tauru
29	39	88.6	241241	12	AC156065	Bos tauru
30	39	88.6	243801	12	AC172361	Bos tauru
31	38	86.4	1020	4	AF508260	Pyrocysti
32	38	86.4	1665	5	HS37S	Pyrocysti
33	38	86.4	3994	13	AY089992	Toxoplasma
34	38	86.4	114177	5	AC090023	Homo sapi
35	38	86.4	115886	12	AC177483	Strongylo
36	38	86.4	151262	5	AC093824	Homo sapi
37	38	86.4	154195	5	AL161773	Human DNA
38	38	86.4	185404	12	AC025409	Homo sapi
39	38	86.4	188518	12	AC166006	Colobus g
40	38	86.4	194158	5	AC021127	Homo sapi
41	38	86.4	194252	6	AC108909	Mus muscu
42	38	86.4	201282	6	AC132450	Mus muscu
43	38	86.4	206760	6	AC122888	Mus muscu
44	38	86.4	217487	12	AC159395	Bos tauru
45	38	86.4	266433	12	AC158817	Bos tauru

ALIGNMENTS

RESULT 1												
BD249732	BD249732											
LOCUS												
DEFINITION												
ACCESSION												
VERSION												
KEYWORDS												
SOURCE												
ORGANISM												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
COMMENT												
	BD249732	Polypeptide.										
	BD249732.1	GI:33059502										
	JP 2002530060-A/2.											
	Mus musculus (house mouse)											
	Mus musculus											
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;											
	Sciurognathi; Muridae; Murinae; Mus...											
	1 (bases 1 to 1281)											
	Carroll,M.W. and Myers,K.A.											
	Polypeptide											
	Patent: JP 2002530060-A 2 17-SEP-2002;											
	OXFORD BIOMEDICA LTD											
	OS Mus musculus (mouse)											
	PN JP 2002530060-A/2											
	PD 17-SEP-2002											
	PF 18-NOV-1999 JP 2000582415											
	PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR											
	30-JUL-1999 GB 9917995.4											
	PI MILES WILLIAM CARROLL,KEVIN ALAN MYERS											
	PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,											
	PC C07K19/00,											
	PC C12N15/00											

```

CC Polypeptide Location/Qualifiers
FH Key 1..1281
FT source /organism="Mus musculus (mouse)"

FEATURES
source
1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 9.13 Length: 1281
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-19 (1-9) x BD249732 (1-1281)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
Db 241 AACCTGCTGGAGGTGCCGCGGATCTA 267

RESULT 2
AX025012 1281 bp DNA linear PAT 15-SEP-2000
LOCUS
DEFINITION Sequence 2 from Patent WO029428.
ACCESSION AX025012
VERSION AX025012.1 GI:10184933
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 Carroll,M.W. and Myers,K.A.
Polypeptide
TITLE
JOURNAL
PATENT: WO 0029428-A 2 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)

FEATURES
source
1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 9.13 Length: 1281
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-19 (1-9) x AX025012 (1-1281)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
Db 241 AACCTGCTGGAGGTGCCGCGGATCTA 267

RESULT 3
AX316087 1281 bp DNA linear PAT 14-DEC-2001
LOCUS
DEFINITION Sequence 2 from Patent EP1160323.
ACCESSION AX316087
VERSION AX316087.1 GI:17899279
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 Carroll,M.W. and Myers,K.A.
Polypeptide
TITLE
JOURNAL
PATENT: EP 1160323-A 2 05-DEC-2001;
OXFORD BIOMEDICA (UK) Limited (GB)

FEATURES
source
1..1281
Location/Qualifiers
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN
Alignment Scores:
Pred. No.: 9.13 Length: 1281
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-19 (1-9) x AX316087 (1-1281)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
Db 241 AACCTGCTGGAGGTGCCGCGGATCTA 267

RESULT 4
AF063939 2333 bp mRNA linear ROD 01-JAN-2000
LOCUS
DEFINITION Rattus norvegicus 5T4 oncofetal antigen homolog (5T4) mRNA,
complete cds.
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE
1 Ninkina,N.N. and Buchman,V.L.
Structure and expression of the rat 5T4 gene
TITLE
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2333)
Buchman,V.L.
Direct Submission
TITLE
Submitted (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK

FEATURES
source
1..2333
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="newborn"
1..2333
/gene="5T4"
1..363
/gene="5T4"
364..1644
/gene="5T4"
/codon_start=1
/protein_id="AAF21770.1"
/db_xref="GI:6650212"
/translation="MPGAGSRGFSAGDGRRLRLARLALVLLGWVSAGAPSSSLPSSSTS
PAAFLASGSAQPPPARCEAEATVKCVNNRLLEVPADLPYPVNNLFTGNQ
MTVLPAGAFARQPLADLAVLNLSGNHLKEVGAGAFHLPGRLRLDLSHNPLNLSAF
TFAGSNVSVSTPDLLEILNHIPTDQKNGSPGVMVAFEGMVAALRSGLALRGL"

gene
5'UTR
CDS

```

HHLELASNHFLYLPDRLLDQLPSLKHLDLRNNLSVLTYSFRNLTHLESLEHEDNAL
KVLHNSITLAEWGLAHVRVFLDNNPWCDYMDVMVSLKETEVPVDPDKARLTCAFPPEK
MNRGLLDLTSSDLDCCDALTPOSLOTSTVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPSLTNLSSNSDV"
1645..2333
/gene="5T4"
2315..2320
/gene="5T4"

3'UTR

polyA_signal

ORIGIN

Alignment Scores:
Pred. No.: 14.7 Length: 2333
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-19 (1-9) x AF063939 (1-2333)

Qy 1 AnLeuLeuGluValProAlaAspLeu 9

Db 604 AACCTGCTGGAGGTGCTGCGGACCTG 630

RESULT 5

BC087011 2361 bp mRNA linear ROD 13-DEC-2004
LOCUS
DEFINITION Rattus norvegicus trophoblast glycoprotein, mRNA (cdna clone
MGC:93332 IMAGE:7193411), complete cds.

ACCESSION BC087011

VERSION BC087011.1 GI:56268819

KEYWORDS MGC.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2361)

Director MGC Project.

Direct Submission

Submitted (02-DEC-2004) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Howard Jacobs

cDNA Library Preparation: Express Genomics

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 186 Row: 0 Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

source

1..2361
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93332 IMAGE:7193411"
/tissue_type="Heart, rat (Brown Norway)"
/clone_lib="NIH MGC_234"
/lab_host="DH10E"
/note="Vector: pExpress1"
1..2361
/gene="tpbg"
/note="synonym: 5T4"
/db_xref="GeneID:83684"
/db_xref="RGD:621453"
384..1644
/gene="tpbg"
/codon_start=1
/product="Tpbg protein"
/protein_id="AAH87011.1"
/db_xref="GI:56268820"
/db_xref="GeneID:83684"
/translation="MPGAGSRGSPAGDRLRLARLALVLGWVSASAPSSSLPSSSTS
PAAFLAGSAQPPPAERCPAECSEAATVKCVNRNLLEVPADLPYVNRNLPTGNQ
MTVLPAGAPARQPLADLAVNLNSGNHLEKVGAGAEHLGLRLDLSHNLPLNSAF
TFAGNSVSTFSPLELILNHIVPPEDQQRQNGSPFGVAFEGMVAALRSGLALRGL
HLELASNHFLYLPDRLLDQLPSLKHLDLRNNLSVLTYSFRNLTHLESLEHEDNAL
KVLHNSITLAEWGLAHVRVFLDNNPWCDYMDVMVSLKETEVPVDPDKARLTCAFPPEK
MNRGLLDLTSSDLDCCDALTPOSLOTSTVFLGIVLALIGAIFLLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPSLTNLSSNSDV"

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.: 14.9 Length: 2361
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-19 (1-9) x BC087011 (1-2361)

Qy 1 AnLeuLeuGluValProAlaAspLeu 9

Db 604 AACCTGCTGGAGGTGCTGCGGACCTG 630

RESULT 6

BC058198

LOCUS

DEFINITION

Mus musculus trophoblast glycoprotein, mRNA (cdna clone MGC:68145

IMAGE:5353871), complete cds.

ACCESSION BC058198

VERSION BC058198.1 GI:34849573

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 2423)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Mallek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 2423)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (15-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tauregeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 123 Row: p Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

source Location/Qualifiers
1..2423
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:68145 IMAGE:5353871"
/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP Mam6"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..2423
/gene="Tpbp"

gene /note="synonym: 574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
402..1682
/gene="Tpbp"

CDS

codon_start=1
product="Tpbp protein"
protein_id="AAH58198.1"
/db_xref="GI:34849574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
translation="MPGASGRPSAGDGLRLRLARLALVLLGWVSASAPSSVPSSTTS
PAAFLASGAQPPAERCAECSEARTVKVNRNLLLEVPADLPYVRNLTGNGQ
MTVLPAGAPARQPLADLEALNSGNHLKEVCAGAFHLGLRLDLSHNPLTNLSAF
APAGSNASVSPLEELTNHIVPPEDQRQNGSPGCVAFEGMVAALRSGLALRGL
TCLEASNHFLPLRDLLAQLPSLRYLDLRNNSLSVLTYSFENLTHTLESHTEDNAL
KVLHNSHTLAEWQGLAHVKVFLDNNPWVCDYMDVMVWLKETEVPDKARLTCAFPFK
MNRNGLLDNNSDLDCAVLPSQTSYVFLGIVLALIGAIFLLVLYLNKRGKIKKWWH
NIRDACRDHMEGYHYRYETADNPRLTNLSSNSDV"
642..1262
/gene="Tpbp"
/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [Function unknown]"
/db_xref="CDD:COG4886"
1299..1415
/gene="Tpbp"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

ORIGIN

Alignment Scores:
Pred. No.: 15.2 Length: 2423
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-19 (1-9) x BC058198 (1-2423)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9
|||||
642 AACCTGCTGGAGGTGCCGCGGATCTA 668

Db

RESULT 7
AX961912 2557 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION Sequence 123 from Patent WO03104277.
ACCESSION AX961912
VERSION AX961912.1 GI:40881322
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.

REFERENCE 1 Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
Stat6 activation gene
Patent: WO 03104277-A 123 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
Location/Qualifiers
1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAP06465.1"
/db_xref="GI:40881323"

FEATURES

source Location/Qualifiers
1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAP06465.1"
/db_xref="GI:40881323"

CDS

translation="MPGASGRPSAGDGLRLRLARLALVLLGWVSASAPSSVPSSTTS
PAAFLASGAQPPAERCAECSEARTVKVNRNLLLEVPADLPYVRNLTGNGQ
MTVLPAGAPARQPLADLEALNSGNHLKEVCAGAFHLGLRLDLSHNPLTNLSAF
APAGSNASVSPLEELTNHIVPPEDQRQNGSPGCVAFEGMVAALRSGLALRGL
TCLEASNHFLPLRDLLAQLPSLRYLDLRNNSLSVLTYSFENLTHTLESHTEDNAL
KVLHNSHTLAEWQGLAHVKVFLDNNPWVCDYMDVMVWLKETEVPDKARLTCAFPFK

MMNRGLLDNSDLDCDAVLPOSLOTSTYVFLGIVLALIGAIFLLVLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNSDV"

ORIGIN

Alignment Scores:

Pred. No.: 15.8 Length: 2557
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-19 (1-9) x AX961912 (1-2557)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9

Db 796 AACCTGCTGGAGTCCCGCGGATCTA 822

RESULT 8

AX961914 AX961914 2557 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 125 from Patent WO03104277.
ACCESSION AX961914
VERSION AX961914.1 GI:40881324

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1 Sugahara, T., Matsuda, A., Honda, G., Muramatsu, S. and Ishizawa, K.
AUTHORS
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 125 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)

FEATURES

source 1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06466.1"
/db_xref="GI:40881325"
/translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSSTS
PADFLASGAOPPPAERCPAAECSEAAATVKVNRNLLEVPADLPVYVRLNLTGNGQ
MTVLPAGAPARQPPPLADLEALNSGNHLKEVCAGAFEHLPGRLRLDLSHNPNTLSAF
VFAGSNASVSPLEELINHIIVPPDQQRNGSFGWVAFEGWVAALRSGLALRGL
TRLEASNHFLFPRDLAQPLSLRYLDLRNNSLSVLTYSFRNLTHLESLEHDNAL
KVLHNSITLAEWQGLAHVKVFLDNNPWCDYCMADVMVWLKETEVPDKARLTCAFPPEK
MMNRGLLDNSDLDCDAVLPOSLOTSTYVFLGIVLALIGAIFLLVLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNSDV"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 15.8 Length: 2557
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-19 (1-9) x AX961914 (1-2557)

Qy 1 AsnLeuLeuGluValProAlaAspLeu 9

Db 796 AACCTGCTGGAGTCCCGCGGATCTA 822

RESULT 9

MMU012160 MMU012160 7942 bp DNA linear ROD 15-APR-2005
LOCUS Mus musculus 5T4 oncofetal trophoblast glycoprotein gene.
DEFINITION

ACCESSION

AJ012160 AJ012160.1 GI:3805948

VERSION 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE

1 King, K.W., Sheppard, F.C., Westwater, C., Stern, P.L. and Myers, K.A.
AUTHORS
TITLE Organisation of the mouse and human 5T4 oncofetal leucine-rich
glycoprotein genes and expression in foetal and adult murine
tissues

JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)

PUBMED

10366710 2 (bases 1 to 7942)

REFERENCE

Myers, K.A.

TITLE Direct Submission

JOURNAL

Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK

FEATURES

source 1..7942
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="Lambda Dash"
3108..3113
/bound_moiety="Spl"
3114..3119
/bound_moiety="Spl"
3124..5779
/gene="5T4"
3124..3151
/gene="5T4"
3152..3450
/gene="5T4"
3451..5779
/gene="5T4"
3779..5059
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA0931.1"
/db_xref="GI:3805949"
/db_xref="GOA:O920L0"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProtKB/T-EMBL:O920L0"
/translation="MPGAGSRGPSAGDGRRLRLARLALVLLGWVSASAPSSVPSSSTS
PADFLASGAOPPPAERCPAAECSEAAATVKVNRNLLEVPADLPVYVRLNLTGNGQ
MTVLPAGAPARQPPPLADLEALNSGNHLKEVCAGAFEHLPGRLRLDLSHNPNTLSAF
VFAGSNASVSPLEELINHIIVPPDQQRNGSFGWVAFEGWVAALRSGLALRGL
TRLEASNHFLFPRDLAQPLSLRYLDLRNNSLSVLTYSFRNLTHLESLEHDNAL
KVLHNSITLAEWQGLAHVKVFLDNNPWCDYCMADVMVWLKETEVPDKARLTCAFPPEK
MMNRGLLDNSDLDCDAVLPOSLOTSTYVFLGIVLALIGAIFLLVLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNSDV"

misc_binding

misc_binding

gene

exon

intron

exon

CDS

sig_peptide

mat_peptide

polyA_signal

polyA_signal

ORIGIN

Alignment Scores:

Pred. No.: 39.2 Length: 7942
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-19 (1-9) x MMU012160 (1-7942)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9

Db 4019 AACCTGCTGGAGGTGCCGCGGATCTA 4045

RESULT 10
 AC158516/c 167046 bp DNA linear ROD 21-JUN-2005
 LOCUS Mus musculus BAC clone RP24-S11A23 from chromosome 9, complete
 DEFINITION

ACCESSION AC158516 AC117768
 VERSION AC158516.2 GI:63025421
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

AUTHORS Adams,S., Cotton,M. and Haglund,K.
 TITLE The sequence of Mus musculus BAC clone RP24-S11A23

JOURNAL Unpublished (2001)
 REFERENCE 2 (bases 1 to 167046)
 AUTHORS Wilson,R.K.

TITLE Direct Submission
 JOURNAL Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 167046)
 AUTHORS Wilson,R.K.

TITLE Direct Submission
 JOURNAL Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 167046)
 AUTHORS Wilson,R.K.

TITLE Direct Submission
 JOURNAL Submitted (21-JUN-2005) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

COMMENT On May 4, 2005 this sequence version replaced gi:61656412.

----- Genome Center
 Center: Washington University Genome Sequencing Center

Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: submissions@wustl.edu

----- Summary Statistics
 Center project name: M_BB0511A23

Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e. phred quality
 >30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone, fosmid clone or direct clone walk sequence.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to obtain the consensus sequence. The
 assembly was confirmed by restriction digest.

This finishing standard has slightly changed from the previous
 Human standard. Specifically, standards for regions of low sequence
 complexity (such as dinucleotide repeats and small unit tandem
 repeats) have been relaxed. These regions are very prevalent in the
 mouse genome, and the return on extended finishing efforts is
 minimal.

If a sequence meets the criteria of the above statement, it needs
 no comments or tags. If the criteria are not met, such as ambiguous
 bases, then the region is duly annotated.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The BAC Library has been constructed by Pieter de Jong and
 coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
 and/or brain genomic DNA. The clone and detailed information can be
 obtained from Pieter de Jong and coworkers at http://www.chori.org

This sequence is the entire insert of the clone.

FEATURES

Location/Qualifiers
 source
 1..167046
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="9"
 /clone_lib="RPCI-24"
 /clone="RP24-S11A23"
 misc_feature
 16685..16712
 /note="Sequence derived from PCR product of genomic DNA"
 unsure
 31565..31779
 /note="Unresolved simple sequence repeat."
 unsure
 46721..46808
 /note="Unresolved simple sequence repeat."
 unsure
 142336..142347
 /note="Sequence derived from one plasmid subclone."

ORIGIN

Alignment Scores:
 Pred. No.: 446 Length: 167046
 Score: 44.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-19 (1-9) x AC158516 (1-167046)

QY 1 AsnLeuLeuGluValProAlaAspLeu 9

Db 110598 AACCTGCTGGAGGTGCCGCGGATCTA 110572

RESULT 11

AC128294/c

LOCUS Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.
 DEFINITION

ACCESSION AC128294

VERSION AC128294.3 GI:25083347

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 210237)
 Muzny,D.,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allien,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

```

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,I., Garza,M.,
Gregregegis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulsegad,H., Lozado,R.J., Lu,X., Ma,J.,
Maneshwari,M., Mahindratne,M., Mamoud,M., Mallory,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Pasakelmech,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-I.,
Puzos,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleciyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 210237)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 210237)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23265004.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZGV
Center clone name: CH230-176H20
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 201781 bases at least Q40
Consensus quality: 203921 bases at least Q20
Consensus quality: 205310 bases at least Q30
Estimated insert size: 205531; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 210237: contig of 210237 bp in length.
FEATURES
source
Location/Qualifiers
1..210237
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-176H20"
1..1142
/note="wgs_end_extension
clone_end:T7"
2177..144799
/note="clone_boundary
clone_end:T7
site:
end_sequence:BH360464"
complement(206062..206961)
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:BH360465"
208907..210237
/note="wgs_end_extension
clone_end:Sp6"
ORIGIN
Alignment Scores:
Pred. No.: 536 Length: 210237
Score: 44.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0
US-10-774-176-19 (1-9) x AC128294 (1-210237)
QY 1 AsnLeuGluValProAlaAepLeu 9
|||||
Db 111278 AACCTGCTGGAGGTGCTGGGACCTG 111252
RESULT 12
AC106962/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4
unordered pieces.
ACCESSION
AC106962
GI:25139469
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 239076)
AUTHORS
Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,

```

Allen, C., Allen, H., Alabrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.B., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebirt, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzuewa, L., Louisedge, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poinexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villanasa, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (14-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GOPI

Center clone name: CH230-87110

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 228642 bases at least Q40

Consensus quality: 232269 bases at least Q30

Consensus quality: 234041 bases at least Q20

Estimated insert size: 231522; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 234710: contig of 234710 bp in length

* 234711 234810: gap of unknown length

* 234811 235924: contig of 1114 bp in length

* 235925 236024: gap of unknown length

* 236025 237314: contig of 1290 bp in length

* 237315 239076: contig of unknown length

* 239076: contig of 1662 bp in length.

FEATURES

source

1. 239076
/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-87110"

234711..234810

/estimated_length=unknown

235925..236024

/estimated_length=unknown

237315..237414

/estimated_length=unknown

ORIGIN

Alignment Scores:

Pred. No.: 594 Length: 239076

Score: 44.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 12 Gaps: 0

US-10-774-176-19 (1-9) x AC106962 (1-239076)

QY 1 AnLeuLeuGluValProAlaAspLeu 9

DB 16489 AACCTGCTGGAGGTGCTCGGACCTG 16463

RESULT 13

AB003707/c

LOCUS

DEFINITION

AB003707

ACCESSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Annelida; Polychaeta; Scolecida; Capitellida;

AB003707

Maldane cristata mRNA for elongation factor-1alpha, partial cds.

1125 bp mRNA linear INV 05-FEB-1999

GI:3063354

elongation factor-1alpha.

Maldane cristata

Maldane cristata

Eukaryota; Metazoa; Annelida; Polychaeta; Scolecida; Capitellida;


```

REFERENCE
AUTHORS      Maldanidae; Maldane.
TITLE        1 (sites)
              Kojima,S.
              Paraphyletic status of Polychaeta suggested by phylogenetic
              analysis based on the amino acid sequences of elongation factor-1
              alpha
JOURNAL      Mol. Phylogenet. Evol. 9 (2), 255-261 (1998)
PUBMED      9562984
AUTHORS      2 (bases 1 to 1125)
              Kojima,S.
TITLE        Direct Submission
JOURNAL      Submitted (07-MAY-1997) Shigeaki Kojima, University of Tokyo, Ocean
              Research Institute; 1-15-1 Minamidai, Nakano-ku, Tokyo 164, Japan
              (E-mail:kojima@trout.ori.u-tokyo.ac.jp, Tel.03-5351-6473,
              Fax:03-5351-6471)
FEATURES
source       Location/Qualifiers
              1..1125
              /organism="Maldane cristata"
              /mol_type="mRNA"
              /db_xref="taxon:73384"
              <!--1125
              /codon_start=1
              /product="elongation factor-1alpha"
              /protein_id="BAA25736.1"
              /db_xref="GI:3063355"
              /translation="QEMGKGSFYAWLVDKLKAERGITIDIALWFKTSKYVTII
              DAPGRDFIKMITGTSQADCAVLVAAGTGEFAGISKNGQTRHALLAYLYGVTKQL
              IVGVNMKDTPTTPYSGPRFDEIKKEGVGYIKKIGNPDTPVPFISGWHGDNMLRKSD
              KMSYNGQWKKVDSKEYKVTLMDALDNIDPPKPTDKALRLPQDVYKIGIGITVP
              VGRVETGLVKPMVVTAPPMTITTEKSVEMHHQALTEAPGDNVGFNINKVSVKDV
              RGNVCGSKNDPPAGTEFEKSVQIILNHPGQIQAGTAPVVDCHTAHIAICRFKELLEKI
              DRRSGKKLEDPNPAHYKSGDACTIVEVPGKSMCVFAFVNPAP"
ORIGIN
Alignment Scores:
Pred. No.:      39.8      Length:      1125
Score:          41.00     Matches:      8
Percent Similarity: 100.0%      Conservative: 1
Best Local Similarity: 88.9%      Mismatches: 0
Query Match:     93.2%      Indels:      0
DB:             13      Gaps:      0

US-10-774-176-19 (1-9) x AB003707 (1-1125)
Oy 1 AsnLeuLeuGluaValProAlaAspLeu 9
|||||...:|||||
Db 405 AATCTTCTGTGTACCGCCGACCTC 379

AC098662      236951 bp      DNA      linear      HTG 10-MAY-2003
Rattus norvegicus clone CH230-11811, *** SEQUENCING IN PROGRESS
***
AC098662      7      GI:30520710
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 236951)
Murny,D,Marie., Metzker,M,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Aisbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

```

```

Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebruggeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kratt,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulsegad,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D.,
Sneed,A., Sodergreen,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D. von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 236951)
Worley,K.C.
Direct Submission
Submitted (28-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 236951)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24942631.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHWC
Center clone name: CH230-11811

```

```

----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 20463 bases at least Q40
Consensus quality: 208139 bases at least Q30
Consensus quality: 210836 bases at least Q20
Estimated insert size: 216195; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 1 contigs. Gaps between the contigs
  * are represented as runs of N. The order of the pieces
  * is believed to be correct as given, however the sizes
  * of the gaps between them are based on estimates that have
  * provided by the submittor.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  * 1 236951: contig of 236951 bp in length.
FEATURES             Location/Qualifiers
     source            1..236951
                        /organism="Rattus norvegicus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10116"
                        /clone="CH230-11811"
     misc_feature      1..1629
                        /note="wgs contig"
     misc_feature      34583..36448
                        /note="wgs contig"
     misc_feature      80470..82578
                        /note="wgs contig"
     misc_feature      225406..226786
                        /note="wgs contig"
     misc_feature      234726..236951
                        /note="wgs contig"

ORIGIN
Alignment Scores:
Pred No.:             2.85e+03      Length:      236951
Score:               41.00          Matches:      8
Percent Similarity:  100.0%          Conservative: 1
Best Local Similarity: 88.9%          Mismatches:   0
Query Match:         93.2%          Indels:       0
DB:                  12             Gaps:         0

US-10-774-176-19 (1-9) x AC098662 (1-236951)
Qy  1 AsnLeuGluValProAlaAspleu 9
    |||:|||||:|||||:|||||:|||||:
Db  149427 AATCTCTGGAGTGCCCGCTGACCTC 149453

RESULT 15
AC125765      244451 bp      DNA      linear      HTG 19-NOV-2002
LOCUS      Rattus norvegicus clone CH230-231L11, WORKING DRAFT SEQUENCE.
ACCESSION      AC125765
VERSION      AC125765.4 GI:25084067
KEYWORDS      HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
Muzny D.Marie., Metzker M.Lee., Abranzon S., Adams C., Alder J.,
Allen C., Allen H., Alsebrooks S., Amin A., Anguiano D.,
Ayalebechi V., Ayodeji A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Blawie K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen R., Chen Y., Chen Z., Chu J.,

```

```

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, S., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Escoto-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hoques, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheua, L., Louleaged, H., Lozado, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., McNeill, T.Z., Meenen, E.,
Mawhiney, S., McLeod, M.P., McNeill, T.Z., Montemayor, J., Moore, S.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willison, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 244451)
Worley, K.C.
Direct Submission
Submitted (30-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 244451)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23907923.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

```

```

----- Project Information
Center project name: G2FE
Center clone name: CH230-231L11
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 237334 bases at least Q40
Consensus quality: 239688 bases at least Q30
Consensus quality: 240866 bases at least Q20
Estimated insert size: 244791; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 244451: contig of 244451 bp in length.
  Location/Qualifiers
    1..244451
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-231L11"
    misc_feature
      1..1665
        /note="wgs_end_extension
        clone_end:Sp6"
    misc_feature
      87432..112936
        /note="clone_boundary
        clone_end:Sp6
        site:_
        end_sequence:BZ111594"
    misc_feature
      complement(240601..241392)
        /note="clone_boundary
        clone_end:T7
        site:_
        end_sequence:BZ111591"
ORIGIN

Alignment Scores:
Pred. No.:      2.92e+03      Length:      244451
Score:          41.00        Matches:      8
Percent Similarity: 100.0%      Conservatives: 1
Best Local Similarity: 88.9%      Mismatches: 0
Query Match:      93.2%         Indels:      0
DB:              12           Gaps:        0

US-10-774-176-19 (1-9) x AC125765 (1-244451)

Qy      1 AsnLeuLeuGluValProAlaAspLeu 9
Db      227838 AATGTCCTGGAGGTGCCGCTGACCTC 227864

```

Search completed: May 27, 2006, 19:35:48
 Job time : 3381.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-18

Perfect score: 50

Sequence: 1 YMADMVAWL 9

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB spool/US10774176/runat_26052006_091441_24976/app query.fasta.1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HRAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs802h
-USER=US10774176 @CGN 1 1 2389 @runat_26052006_091441_24976 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	50	100.0	1281	3	AAA27059	Aaa27059 Mouse 574
2	50	100.0	2557	12	ADI26160	Adi26160 Human cDN
3	50	100.0	2557	12	ADI26158	Adi26158 Human cDN

4	45	90.0	1263	4	AAF89736	Aaf89736 Nucleotid
5	45	90.0	1263	6	ABK87174	Abk87174 cDNA enco
6	41	82.0	299	10	ACD93536	AcD93536 Human col
7	41	82.0	453	5	AAS87174	Aas87174 DNA enco
8	41	82.0	475	13	ADU11677	AdU11677 Solid tum
9	41	82.0	927	6	ABT07721	Abt07721 Breast ca
10	41	82.0	927	8	ABX76333	Abx76333 Lung canc
11	41	82.0	927	10	ADB80503	AdB80503 Ovarian c
12	41	82.0	927	11	ADN38723	Adn38723 Cancer/an
13	41	82.0	973	8	AAD56198	Aad56198 Human LRR
14	41	82.0	1156	6	ABV99349	Abv99349 Human NOV
15	41	82.0	1260	6	ABK87175	Abk87175 cDNA enco
16	41	82.0	1260	10	ADB97513	AdB97513 Feline 5T
17	41	82.0	1260	10	ADB97452	AdB97452 DNA enco
18	41	82.0	1263	3	AAA27058	Aaa27058 Human 5T4
19	41	82.0	1331	8	AAD56199	Aad56199 Human LRR
20	41	82.0	2020	10	ADJ56299	Adj56299 Human cDN
21	41	82.0	2053	8	ACC51052	Acc51052 Human bla
22	41	82.0	2053	8	ABX76332	Abx76332 Lung canc
23	41	82.0	2053	8	AAD56197	Aad56197 Human LRR
24	41	82.0	2053	8	AAD56200	Aad56200 Human LRR
25	41	82.0	2053	11	ADN38721	Adn38721 Cancer/an
26	41	82.0	2053	12	ADL06473	Adl06473 Human tum
27	41	82.0	2053	12	ADN03961	Adn03961 Antipsori
28	41	82.0	2053	13	ADR25444	Adr25444 Breast ca
29	41	82.0	2053	13	ACN38510	Acn38510 Tumour-as
30	41	82.0	2053	13	ADV35098	Adv35098 Human cDN
31	41	82.0	2053	14	AED17761	Aed17761 Fibrotic
32	41	82.0	2338	5	AAS87175	Aas87175 DNA enco
33	41	82.0	2359	4	AAK94253	Aak94253 Human ful
34	41	82.0	2359	12	ADL30831	Adl30831 Full leng
35	41	82.0	2361	4	AAK94254	Aak94254 Human ful
36	41	82.0	2361	12	ADI26162	Adi26162 Human cDN
37	41	82.0	2361	12	ADL30833	Adl30833 Full leng
38	40	80.0	891	13	ADS57375	Ads57375 Bacterial
39	40	80.0	1032	13	ADS57236	Ads57236 Bacterial
40	40	80.0	1143	13	ADT42735	Adt42735 Bacterial
41	40	80.0	2391	4	AAF77826	Aaf77826 SfUCPa de
42	39	78.0	882	2	AAV61796	Aav61796 Subcloned
43	39	78.0	1032	13	ADS64173	Ads64173 Bacterial
44	39	78.0	1032	13	ADT46017	Adt46017 Bacterial
45	39	78.0	1032	13	ADS63789	Ads63789 Bacterial

ALIGNMENTS

RESULT 1
AAA27059
ID AAA27059 standard; DNA; 1281 BP.
XX
AC
AAA27059;
XX
DT 22-AUG-2000 (first entry)
XX
DE Mouse 574 tumour-associated antigen gene.
XX
KW Mouse; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
KW ds.
XX
OS Mus musculus.
XX
PN WO200029428-A2.
XX
PD 25-MAY-2000.
XX
PF 18-NOV-1999; 99WO-GB003859.
XX
PR 18-NOV-1998; 98GB-00025103.
PR 27-JAN-1999; 99GB-00001739.
PR 30-JUL-1999; 99GB-00017995.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Carroll MW, Myers KA;
 XX WPI; 2000-387735/33.
 XX
 PT Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.
 XX
 PS Example 2; Page 78; 79pp; English.
 XX
 CC The present sequence encodes the mouse 5T4 tumour-associated antigen
 CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
 CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The 5T4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject. The
 CC present sequence appears in GenBank at accession number AJ012160
 XX
 SQ Sequence 1281 BP; 246 A; 410 C; 354 G; 271 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.54 Length: 1281
 Score: 50.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-18 (1-9) x AAA27059 (1-1281)
 Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
 Db 919 TACATGGCTGACATGGTGGCTTGGCTT 945
 RESULT 2
 ADI26160
 ID ADI26160 standard; cDNA; 2557 BP.
 XX AC ADI26160;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #63.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 PF 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00137505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 PA (ASAH) ASahi KASEI KK.
 XX

PI Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 XX WPI; 2004-122214/12.
 DR P-PSDB; ADI26161.
 XX
 PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 PS Claim 4; SEQ ID NO 125; 1368pp; English.
 XX
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infections disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 2557 BP; 513 A; 729 C; 696 G; 619 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9.6 Length: 2557
 Score: 50.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0
 US-10-774-176-18 (1-9) x ADI26160 (1-2557)
 Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
 Db 1474 TACATGGCTGACATGGTGGCTTGGCTT 1500
 RESULT 3
 ADI26158
 ID ADI26158 standard; cDNA; 2557 BP.
 XX AC ADI26158;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human cDNA encoding protein that promotes STAT6 activation #62.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX

DT 07-OCT-2002 (first entry)
 XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
 DE
 XX
 XX Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytotstatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX
 XX Canis sp.
 XX
 XX Key Location/Qualifiers
 FH 1..1263
 FT CDS /*tag= a
 FT /product= "5T4 protein"
 XX
 XX WO200238612-A2.
 XX
 XX 16-MAY-2002.
 XX
 XX 13-NOV-2001; 2001WO-GB005004.
 XX
 XX 13-NOV-2000; 2000WO-GB004317.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 XX Myers K, Drury N, Carroll M;
 XX
 XX WPI; 2002-557449/59.
 DR P-PSDB; AAU98693.
 DR
 XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.
 XX
 XX Claim 1; Page 67; 68pp; English.
 PS
 XX The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes canine 5T4 protein
 XX
 XX Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 41.7 Length: 1263
 Score: 45.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.8% Mismatches: 0
 Query Match: 90.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-18 (1-9) x ABK87174 (1-1263)
 Qy 1 TyrMetAlaSepMetValAlaTrpLeu 9
 Db 901 CACATGGCAGACATGGTGGCTGGCTC 927
 RESULT 6
 ACD93536

ID ACD93536 standard; cDNA; 299 BP.
 XX ACD93536;
 AC
 XX 23-SEP-2003 (first entry)
 DT
 XX Human colon cancer cell expressed cDNA #1948.
 DE
 XX Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KW agriculture; food crop genome; resistance gene; retrovirus;
 KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX US2002155438-A1.
 PN
 XX 24-OCT-2002.
 PD
 XX 27-SEP-1999; 99US-00406117.
 PF
 XX 20-NOV-1998; 98US-00196716.
 PR
 XX (SIMP/) SIMPSON A J G.
 PA (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 PA
 XX Simpson AJG, Neto ED, Brentani RR;
 PI WPI; 2003-182626/18.
 XX
 DR Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 PT
 XX Example 9; Page 302; 959pp; English.
 PS
 XX The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (I) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a foetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library

XX Sequence 299 BP; 75 A; 84 C; 78 G; 62 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 52.2 Length: 299
 Score: 41.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 82.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-18 (1-9) x ACD93536 (1-299)

Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
 Db 195 CACATGGCAGACATGTCACCTGGCTC 221

RESULT 7
 AAS87174
 ID AAS87174 standard; cDNA; 453 BP.
 AC AAS87174;
 XX
 XX
 XX 13-FEB-2002 (first entry)
 XX
 XX DNA encoding novel human diagnostic protein #22978.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 FN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 FF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 PR 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 FI
 XX WPI; 2001-639362/73.
 DR
 DR P-PSDB; ASG22987.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 XX Claim 1; SEQ ID NO 22978; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 453 BP; 108 A; 111 C; 113 G; 121 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 81.9 Length: 453
 Score: 41.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 82.0% Indels: 0
 DB: 5 Gaps: 0

US-10-774-176-18 (1-9) x AAS87174 (1-453)

Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
 Db 88 CACATGGCAGACATGTCACCTGGCTC 114

RESULT 8
 ADU11677
 ID ADU11677 standard; DNA; 475 BP.
 XX
 XX ADU11677;
 AC
 XX
 XX 27-JAN-2005 (first entry)
 DT
 XX Solid tumour prognosis gene seqid 2116.
 XX Cytostatic; gene therapy; expression profile; solid tumour;
 KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
 XX Unidentified.
 OS
 XX WO2004097052-A2.
 FN
 XX 11-NOV-2004.
 PD
 XX 29-APR-2004; 2004WO-US013587.
 XX
 XX 29-APR-2003; 2003US-0466067P.
 PR
 PR 23-JAN-2004; 2004US-0538246P.
 XX
 XX (AMHP) WYETH.
 PA
 PA (STRA/) STRAHS A.
 XX
 XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
 FI Immerman F, Dörner AJ;
 XX
 XX WPI; 2004-804779/79.

A method, useful for prognosing and treating solid tumor, comprises comparing an expression profile of a gene expressed in peripheral blood mononuclear cells to a reference expression profile of a gene.
 Disclosure; Page; 111pp; English.

The invention describes a method comprising comparing an expression profile of at least one gene in a peripheral blood sample of a patient to at least one reference expression profile of the at least one gene, where the patient has a solid tumour, and each of the gene is differentially expressed in peripheral blood mononuclear cells (PBMCs) of a first class of patients as compared to PBMCs of a second class of patients, where both the first and second classes of patients have the solid tumour, and each of the first and second classes is a subcluster formed by an unsupervised clustering analysis of gene expression profiles in PBMCs of a population of patients who have the solid tumour, and where the majority of the first class of patients has a first clinical outcome, and the majority of the second class of patients has a second clinical outcome. Also described are: a system comprising (i) a memory or a storage medium including data that represent an expression profile of at least one gene in a peripheral blood sample of a patient who has a solid

XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX Claim 22; Page 336; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 178 Length: 927
 Score: 41.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 82.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-18 (1-9) x ABX76333 (1-927)

Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
 ::::::::::::::|
 Db 559 CACATGGCAGACATGGTGACCTGGCTC 585

RESULT 11
 ADB80503
 ID ADB80503 standard; DNA; 927 BP.

XX ADB80503;
 XX 04-DEC-2003 (first entry)
 XX Ovarian cancer-associated transcript #34.
 XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection; ds; gene.

XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 1..927
 FT /*tag= a

XX WO2002102235-A2.
 XX 27-DEC-2002.
 XX 18-JUN-2002; 2002WO-US019297.
 XX 18-JUN-2001; 2001US-0299234P.
 XX 27-AUG-2001; 2001US-0315287P.
 XX 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

DR P-PSDB; ADB80504.

XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

XX Claim 10; Page 297; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.

XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 178 Length: 927
 Score: 41.00 Matches: 7
 Percent Similarity: 88.9% Conservative: 1
 Best Local Similarity: 77.8% Mismatches: 1
 Query Match: 82.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-18 (1-9) x ADB80503 (1-927)

Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
 ::::::::::::::|
 Db 559 CACATGGCAGACATGGTGACCTGGCTC 585

RESULT 12
 ADB38723

ID ADB38723 standard; cDNA; 927 BP.

XX ADB38723;

DT 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

OS WO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

```

PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 23-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
XX (BOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zioltnik A;
XX
XX WPI: 2003-468649/44.
DR P-PSDB; ADN38724.
XX
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX Claim 8; SEQ ID NO 41; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 178 Length: 927
Score: 41.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 82.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-18 (1-9) x ADN38723 (1-927)

QY 1 TyrMetAlaAspMetValAlaTrpLeu 9
ID ::|||:|||:|||:|||
Db 559 CACATGGCAGACATGGTGACCTGGCTC 585

RESULT 13
AAD56198
ID AAD56198 standard; DNA; 973 BP.
XX
XX AC AAD56198;
XX

```

KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX
OS Homo sapiens.

XX
XX WO200272771-A2.

XX
XX 19-SEP-2002.

XX
XX 08-MAR-2002; 2002WO-US007288.

XX
XX 08-MAR-2001; 2001US-0274101P.

XX
XX 08-MAR-2001; 2001US-0274194P.

XX
XX 08-MAR-2001; 2001US-0274281P.

XX
XX 08-MAR-2001; 2001US-0274322P.

XX
XX 09-MAR-2001; 2001US-0274849P.

XX
XX 12-MAR-2001; 2001US-0275233P.

XX
XX 13-MAR-2001; 2001US-0275578P.

XX
XX 13-MAR-2001; 2001US-0275579P.

XX
XX 13-MAR-2001; 2001US-0275601P.

XX
XX 14-MAR-2001; 2001US-0276000P.

XX
XX 16-MAR-2001; 2001US-0276776P.

XX
XX 19-MAR-2001; 2001US-0276994P.

XX
XX 20-MAR-2001; 2001US-0277239P.

XX
XX 20-MAR-2001; 2001US-0277321P.

XX
XX 20-MAR-2001; 2001US-0277327P.

XX
XX 20-MAR-2001; 2001US-0277338P.

XX
XX 21-MAR-2001; 2001US-0277791P.

XX
XX 22-MAR-2001; 2001US-0277833P.

XX
XX 23-MAR-2001; 2001US-0278152P.

XX
XX 26-MAR-2001; 2001US-0278894P.

XX
XX 27-MAR-2001; 2001US-0278999P.

XX
XX 27-MAR-2001; 2001US-0279036P.

XX
XX 28-MAR-2001; 2001US-0279344P.

XX
XX 30-MAR-2001; 2001US-0279995P.

XX
XX 30-MAR-2001; 2001US-0280233P.

XX
XX 02-APR-2001; 2001US-0280802P.

XX
XX 02-APR-2001; 2001US-0280822P.

XX
XX 04-APR-2001; 2001US-0280900P.

XX
XX 13-APR-2001; 2001US-0281194P.

XX
XX 30-APR-2001; 2001US-0283675P.

XX
XX 02-MAY-2001; 2001US-0288066P.

XX
XX 03-MAY-2001; 2001US-0288342P.

XX
XX 03-MAY-2001; 2001US-0288528P.

XX
XX 15-MAY-2001; 2001US-0291190P.

XX
XX 16-MAY-2001; 2001US-0291099P.

PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX
XX (CURA-) CURAGEN CORP.

XX
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;

PI Boldog FL, Li L, Zernhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;

PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;

PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CS;

PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VV, Pochart PR;

XX
XX Zhong M;

XX
XX WPI; 2002-732824/79.

DR P-PSDB; ABP70071.

XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,

PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,

PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic

PT disorders, and asthma.

XX
XX Claim 16; Page 114-115; 619pp; English.

XX
XX The present invention relates to new isolated proteins (NOVX) and their

CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is

CC any number from 1 to 48. The NOVX proteins and coding sequences are

CC useful in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, preferably a NOVX-associated disorder.

CC The NOVX coding sequences and proteins are useful for treating, diabetes,

CC preventing or diagnosing diseases such as metabolic disorders, diabetes,

CC obesity, infectious disease, anorexia, cancer-associated cachexia,

CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's

CC disease, immune disorders, haematopoietic disorders, cardiovascular

CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic

CC disturbances associated with obesity, metabolic syndrome X or wasting

CC disorders associated with chronic diseases or various cancers. The NOVX

CC coding sequences and proteins may also be used as targets for the

CC identification of small molecules that modulate or inhibit e.g.

CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,

CC wound healing and angiogenesis, in gene therapy, in generation of

CC antibodies that bind immunospecifically to NOVX substances for use in

CC therapeutic or diagnostic methods

XX
XX SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 226 Length: 1156

Score: 41.00 Matches: 7

Percent Similarity: 88.9% Conservative: 1

Best Local Similarity: 77.8% Mismatches: 1

Query Match: 82.0% Indels: 0

DB: Gaps: 0

US-10-774-176-18 (1-9) x ABV99349 (1-1156)

QY 1 TyrMetAlaaspMetValAlaTrpLeu 9

Db 790 CACATGGCAGCATGTGACCTGGCTC 816

RESULT 15

ABK87175

ID ABK87175 standard; cDNA; 1260 BP.

XX

AC ABK87175;

XX

DT 07-OCT-2002 (first entry)

XX

DE cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.

XX

KW Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;

KW cell proliferative disorder; infection; inflammatory condition;

KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
XW foetal abnormality; foetal sex determination; gene; ss.

XX Felis sp.

XX OS Location/Qualifiers

XX FH 1..1260

FT CDS /*tag= a

FT FT /product= "5T4 protein"

XX WO200238612-A2.

PN 16-MAY-2002.

XX 13-NOV-2001; 2001WO-GB005004.

XX 13-NOV-2000; 2000WO-GB004317.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Myers K, Drury N, Carroll M;

XX WPI; 2002-557449/59.

DR P-PSDB; AAU98694.

XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.

XX Claim 4; Page 68; 69pp; English.

XX The present invention relates to the isolation of canine and feline
CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 5T4 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 5T4
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus(es).
CC The present sequence encodes feline 5T4 protein

XX SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 248 Length: 1260
Score: 41.00 Matches: 7
Percent Similarity: 88.9% Conservative: 1
Best Local Similarity: 77.8% Mismatches: 1
Query Match: 82.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-18 (1-9) x ABK87175 (1-1260)

Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9

Db 898 CACATGGTGGACATGGTGGCTGCTC 924

Search completed: May 27, 2006, 10:38:49
Job time : 379.5 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds
(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-18

Perfect score: 50

Sequence: 1 YWADNVWL 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10774176/runat_26052006_091443_24987/app_query.fasta_1
-DB=GenEmbl -QFWT=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss04
-USER=US10774176 @CGN 1 1 7534 @runat_26052006_091443_24987 -NCPU=6 -ICPU=3
-NO_MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sts.*
8: gb_sy.*
9: gb_un.*
10: gb_vi.*
11: gb_ov.*
12: gb_htg.*
13: gb_in.*
14: gb_on.*
15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	50	100.0	1281	2	BD249732	Polypeptide
2	50	100.0	1281	2	AX025012	Sequence
3	50	100.0	1281	2	AX316087	Sequence

4	50	100.0	2423	6	BC058198	Mus muscu	
5	50	100.0	2557	2	AX961912	Sequence	
6	50	100.0	2557	2	AX961914	Sequence	
7	50	100.0	7942	6	MMU012160	Mus muscu	
c	8	50	100.0	167046	6	AC158516	Mus muscu
	9	47	94.0	2333	6	AF063939	Rattus no
	10	47	94.0	2361	6	BC087011	Rattus no
	11	47	94.0	210237	12	AC128294	Rattus no
c	12	47	94.0	239076	12	AC106962	Rattus no
	13	45	90.0	1263	2	AX149553	Sequence
	14	45	90.0	1263	2	AX467371	Sequence
c	15	42	84.0	75778	15	AY328003	Symbiont
	16	42	84.0	110000	4	AP008211	Continuation (210
	17	42	84.0	110000	15	CP000082	Continuation (22 o
	18	42	84.0	127601	12	DQ157839	Triticum
	19	42	84.0	137650	4	AC130600	Oryza sat
c	20	41	82.0	421	7	HSPA32B9	Z94208 H.sapiens f
	21	41	82.0	475	2	CQ920916	Sequence
	22	41	82.0	927	2	AX829164	Sequence
	23	41	82.0	1156	2	DD161112	Novel Ant
	24	41	82.0	1260	2	AX467373	Sequence
	25	41	82.0	1260	2	AX821533	Sequence
	26	41	82.0	1260	2	AX821548	Sequence
	27	41	82.0	1263	2	BD249731	Polypepti
	28	41	82.0	1263	2	AX025011	Sequence
	29	41	82.0	1263	2	AX316086	Sequence
	30	41	82.0	2053	2	CQ731678	Sequence
	31	41	82.0	2053	2	DD174290	METHODS O
	32	41	82.0	2053	5	HS5T4OA	Z29083 Homo sapien
	33	41	82.0	2359	2	BD127282	Primer fo
	34	41	82.0	2359	2	CQ782724	Sequence
	35	41	82.0	2359	5	AK074786	Homo sapi
	36	41	82.0	2361	2	AX961916	Sequence
	37	41	82.0	2361	2	BD127283	Primer fo
	38	41	82.0	2361	2	CQ782726	Sequence
	39	41	82.0	2361	5	AK074790	Homo sapi
	40	41	82.0	2379	5	BC037161	Homo sapi
	41	41	82.0	2714	5	AB168308	Macaca fa
	42	41	82.0	5551	5	HSA012159	Homo sapi
	43	41	82.0	121909	5	HSJ492P14	Human DNA
c	44	41	82.0	182633	6	AL671215	Mouse DNA
c	45	41	82.0	203593	6	AL663053	Mouse DNA

ALIGNMENTS

RESULT 1	BD249732	1281 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD249732				
DEFINITION	Polypeptide.				
ACCESSION	BD249732				
VERSION	BD249732.1	GI:33059502			
KEYWORDS	JP 2002530060-A/2.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;				
	Sciurognathi; Muridea; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1281)				
AUTHORS	Carroll, M.W. and Myers, K.A.				
TITLE	Polypeptide				
JOURNAL	Patent: JP 2002530060-A 2 17-SEP-2002;				
	OXFORD BIOMEDICA LTD				
COMMENT	OS Mus musculus (mouse)				
	PN JP 2002530060-A/2				
	PD 17-SEP-2002				
	PF 18-NOV-1999 JP 2000582415				
	PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR				
	10-JUL-1999 GB 9917995.4				
FI	MILES WILLIAM CARROLL, KEVIN ALAN MYERS				
PC	C12N15/09, A61K39/00, A61P35/00, C07K7/06, C07K14/065,				
PC	C07K19/00,				
PC	C12N15/00				

```

CC      Polypeptide
FH      Key
FT      source
FT      Location/Qualifiers
FEATURES             source
    1..1281
     /organism="Mus musculus"
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.:          0.325      Length:      1281
Score:              50.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches:  0
Query Match:        100.0%      Indels:       0
DB:                  2          Gaps:         0
US-10-774-176-18 (1-9) x BD249732 (1-1281)
QY      1 TyrMetAlaAspMetValAlaTrpLeu 9
Db      919 TACATGGCTGACATGGTGGCTTGGCTT 945
RESULT 2
AX025012      1281 bp      DNA      linear      PAT 15-SEP-2000
LOCUS
DEFINITION
Sequence 2 from Patent WO029428.
ACCESSION
AX025012
VERSION
AX025012.1 GI:10184933
KEYWORDS
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 Carroll, M.W. and Myers, K.A.
Polypeptide
Patent: WO 0029428-A 2 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES             source
    1..1281
     /organism="Mus musculus"
     /mol_type="unassigned DNA"
     /db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.:          0.325      Length:      1281
Score:              50.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches:  0
Query Match:        100.0%      Indels:       0
DB:                  2          Gaps:         0
US-10-774-176-18 (1-9) x AX025012 (1-1281)
QY      1 TyrMetAlaAspMetValAlaTrpLeu 9
Db      919 TACATGGCTGACATGGTGGCTTGGCTT 945
RESULT 3
AX316087      1281 bp      DNA      linear      PAT 14-DEC-2001
LOCUS
DEFINITION
Sequence 2 from Patent EP1160323.
ACCESSION
AX316087
VERSION
AX316087.1 GI:17899279
KEYWORDS
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Pahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
TITLE
JOURNAL
PUBMED
2 (bases 1 to 2423)
AUTHORS
Strausberg, R.
Direct Submission
Submitted (15-SEP-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.nci.nih.gov

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 Carroll, M.W. and Myers, K.A.
5t4 tumour-associated antigen for use in tumour immunotherapy
Patent: EP 1160323-A 2 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
Location/Qualifiers
    1..1281
     /organism="Mus musculus"
     /mol_type="unassigned DNA"
     /db_xref="taxon:10090"
ORIGIN
Alignment Scores:
Pred. No.:          0.325      Length:      1281
Score:              50.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches:  0
Query Match:        100.0%      Indels:       0
DB:                  2          Gaps:         0
US-10-774-176-18 (1-9) x AX316087 (1-1281)
QY      1 TyrMetAlaAspMetValAlaTrpLeu 9
Db      919 TACATGGCTGACATGGTGGCTTGGCTT 945
RESULT 4
BC058198      2423 bp      mRNA      linear      ROD 21-OCT-2003
LOCUS
DEFINITION
Mus musculus trophoblast glycoprotein, mRNA (cDNA clone MGC:68145
IMAGE:5353871), complete cds.
ACCESSION
BC058198
VERSION
BC058198.1 GI:34849573
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2423)
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Pahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
TITLE
JOURNAL
PUBMED
2 (bases 1 to 2423)
AUTHORS
Strausberg, R.
Direct Submission
Submitted (15-SEP-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.nci.nih.gov

```

COMMENT

Contact: MGC help desk
Email: gcgbs@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,G., Pearson,R., Stantropop,S., Thomas,P.J., Touchman,J.W.,
Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 123 Row: p Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6755854.

FEATURES

source

1..2423
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:68145 IMAGE:5353871"
/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene

1..2423
/gene="Tpbpg"
/note="synonym: 574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
402..1682
/gene="Tpbpg"
/codon_start=1
/product="Tpbpg protein"
/protein_id="AAH58198.1"
/db_xref="GI:34849574"
/db_xref="GeneID:21983"
/db_xref="MGI:1341264"
/translation="MPGAGSRGPSAGDGRRLRLRLALVLLGWVSAPSSSPSSSTS
PAAFLAGSAQPPPAERCAECSEAAATVKVNRNLLVLPADLPVYRNPLFTGNQ
MTVLPAQAPARQPPPLADLEALNLSGNHLKEVCAGAFHLPGLRLDLSHNPLTNSAF
APAGSNASVAPSPLLEELIINHIVPPEDQRONGSPEGMVAFEGMVAALRSGLAGRL
TCLEASNHFLPLRDLAQPLSLRYLDLRNNSLSVLTYSFARNLTLSLHLEDNAL
KVLHNSTLAEMOGLAHVFLDNNPWDCYCNMADVMWLKETEVPDPKARLTCAPEK
MRNGLDLNSSLDCDAVLPGSLQTSYVFLGIVLALIGAILFLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"

CDS

642..1262
/gene="Tpbpg"
/note="COG4886; Region: COG4886, Leucine-rich repeat (LRR)
protein [Function unknown]"
/db_xref="CDD:COG4886"
1299..1415
/gene="Tpbpg"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

misc_feature

1299..1415
/gene="Tpbpg"
/note="LRRCT; Region: Leucine rich repeat C-terminal
domain"
/db_xref="CDD:smart00082"

misc_feature

0.666 Length: 2423
50.00 Matches: 9
100.0% Conservative: 0

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-18 (1-9) x BC058198 (1-2423)
QY 1 TyrMetAlaAspMetValAlaTrpLeu 9
Db 1320 TACATGGCTGACATGCTGGCTTGCTT 1346
RESULT 5
AX961912
LOCUS AX961912
DEFINITION Sequence 123 from Patent WO03104277.
ACCESSION AX961912
VERSION AX961912.1 GI:40881322
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
TITLE Stat6 activation gene.
JOURNAL Patent: WO 03104277-A 123 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06465.1"
/db_xref="GI:40881323"
/translation="MEGAGSRGPSAGDGRRLRLRLALVLLGWVSAPSSSPSSSTS
PAAFLAGSAQPPPAERCAECSEAAATVKVNRNLLVLPADLPVYRNPLFTGNQ
MTVLPAQAPARQPPPLADLEALNLSGNHLKEVCAGAFHLPGLRLDLSHNPLTNSAF
VFGSNASVAPSPLLEELIINHIVPPEDQRONGSPEGMVAFEGMVAALRSGLAGRL
TCLEASNHFLPLRDLAQPLSLRYLDLRNNSLSVLTYSFARNLTLSLHLEDNAL
KVLHNSTLAEMOGLAHVFLDNNPWDCYCNMADVMWLKETEVPDPKARLTCAPEK
MRNGLDLNSSLDCDAVLPGSLQTSYVFLGIVLALIGAILFLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"
ORIGIN
Alignment Scores:
Pred. No.: 0.707 Length: 2557
Score: 50.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-18 (1-9) x AX961912 (1-2557)
QY 1 TyrMetAlaAspMetValAlaTrpLeu 9
Db 1474 TACATGGCTGACATGCTGGCTTGCTT 1500
RESULT 6
AX961914
LOCUS AX961914
DEFINITION Sequence 125 from Patent WO03104277.
ACCESSION AX961914
VERSION AX961914.1 GI:40881324
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.


```

REFERENCE
1 Sugahara,T., Mateuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
AUTHORS Stat6 activation gene
TITLE
JOURNAL Patent: WO 03104277-A 125 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source Location/Qualifiers
1..2557
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
556..1836
/feature="unnamed protein product"
/codon_start=1
/db_xref="GI:40881325"
/translation="MPGAGSGPGAGDGRRLRLALVLLGWVSASAPSSVPSSSTS
PADFLASGSAQPPPAECSEARCTKVCNRLLEVPADLPVYVRLNLTGNQ
MTVLPAGAFARQPPADLEALNSGNHLKVCAGAFHLPGLRLDLSHNPLTNLSAF
VFGNSASVSPLEELIHLNHI VPPEDQRQNGSFGMVAFEGMVAALRSGLALRGL
TRLEASNHFLPLRDLAQLPSLYLDLRNNSLSVLTYSFRLNTHLSLHLEDNAL
KVLHNSLAEWQGLAHVKVFLDNNPWCDYMDVMWMLKETEVPDKARLTCAPEK
MRNRLGLDLSDDCDVLPQSLQTSYVFLGIVLALIGAILFLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"
ORIGIN
Alignment Scores:
Pred. No.: 0.707 Length: 2557
Score: 50.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-18 (1-9) x AX961914 (1-2557)
Qy 1 TyrMetAlaAspMetValAlaTrioLeu 9
Db 1474 TACATGGCTGACATGGTGGCTTGGCTT 1500
RESULT 7
MMU012160 7942 bp DNA linear ROD 15-APR-2005
LOCUS Mus musculus 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION AJ012160
VERSION AJ012160.1 GI:3805948
KEYWORDS 5T4 gene; 5T4 oncofetal trophoblast glycoprotein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 King,K.W., Sheppard,F.C., Westwater,C., Stern,P.L. and Myers,K.A.
AUTHORS Organisation of the mouse and human 5T4 oncofetal leucine-rich
TITLE Glycoprotein genes and expression in foetal and adult murine
tissues
JOURNAL Biochim. Biophys. Acta 1445 (3), 257-270 (1999)
PUBMED 10366710
REFERENCE
2 (bases 1 to 7942)
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Myers K.A., CRC Immunology Group, Paterson
Institute for Cancer Research, Christie Hospital, Wilmslow Road,
Manchester, M20 9BX, UK
FEATURES
source Location/Qualifiers
1..7942
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="lambda Dash"
3108..3113
/misc_binding /bound_moiety="Spl"

```

```

misc_binding 3114..3119
/bound_moiety="Spl"
gene 3124..5779
/gene="5T4"
exon 3124..3151
/gene="5T4"
intron 3152..3450
/gene="5T4"
exon 3451..5779
/gene="5T4"
CDS 3779..5059
/gene="5T4"
/codon_start=1
/product="5T4 oncofetal trophoblast glycoprotein"
/protein_id="CAA09931.1"
/db_xref="GI:3805949"
/db_xref="GOA:Q9Z0L0"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR001611"
/db_xref="InterPro:IPR003591"
/db_xref="MGI:1341264"
/db_xref="UniProtKB/TrEMBL:Q9Z0L0"
/translation="MPGAGSGPGAGDGRRLRLALVLLGWVSASAPSSVPSSSTS
PADFLASGSAQPPPAECSEARCTKVCNRLLEVPADLPVYVRLNLTGNQ
MTVLPAGAFARQPPADLEALNSGNHLKVCAGAFHLPGLRLDLSHNPLTNLSAF
VFGNSASVSPLEELIHLNHI VPPEDQRQNGSFGMVAFEGMVAALRSGLALRGL
TRLEASNHFLPLRDLAQLPSLYLDLRNNSLSVLTYSFRLNTHLSLHLEDNAL
KVLHNSLAEWQGLAHVKVFLDNNPWCDYMDVMWMLKETEVPDKARLTCAPEK
MRNRLGLDLSDDCDVLPQSLQTSYVFLGIVLALIGAILFLVLYLNKRGIKKWMH
NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"
3779..3865
/gene="5T4"
3866..5056
/gene="5T4"
/polya_signal /product="5T4 oncofetal trophoblast glycoprotein"
5713..5718
/gene="5T4"
5759..5764
/gene="5T4"
ORIGIN
Alignment Scores:
Pred. No.: 2.54 Length: 7942
Score: 50.00 Matches: 9
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-18 (1-9) x MMU012160 (1-7942)
Qy 1 TyrMetAlaAspMetValAlaTrioLeu 9
Db 4697 TACATGGCTGACATGGTGGCTTGGCTT 4723
RESULT 8
AC158516 167046 bp DNA linear ROD 21-JUN-2005
LOCUS Mus musculus BAC clone Rp24-511A23 from chromosome 9, complete
DEFINITION sequence.
ACCESSION AC158516 AC117768
VERSION AC158516.2 GI:63025421
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 167046)
AUTHORS Adams,S., Cotton,M. and Haglund,K.
TITLE The sequence of Mus musculus BAC clone Rp24-511A23
JOURNAL Unpublished (2001)

```

REFERENCE AUTHORS TITLE JOURNAL	2 (bases 1 to 167046) Wilson,R.K. Direct Submission Submitted (19-MAR-2005) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE AUTHORS JOURNAL	3 (bases 1 to 167046) Wilson,R.K. Direct Submission Submitted (04-MAY-2005) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE AUTHORS TITLE JOURNAL	4 (bases 1 to 167046) Wilson,R.K. Direct Submission Submitted (21-JUN-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT	On May 4, 2005 this sequence version replaced gi:161556412. ----- Genome Center Center: Washington University Genome Sequencing Center Center code: WUGSC Web site: http://genome.wustl.edu Contact: submissions@watson.wustl.edu ----- Summary Statistics Center project name: M_BB0511A23 Drafting center: WIBR

NOTICE:

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e. phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence. The assembly was confirmed by restriction digest. This finishing standard has slightly changed from the previous Human standard. Specifically, standards for regions of low sequence complexity (such as dinucleotide repeats and small unit tandem repeats) have been relaxed. These regions are very prevalent in the mouse genome, and the return on extended finishing efforts is minimal.

MAPPING INFORMATION.

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

This sequence is the entire insert of the clone.

FEATURES	source	Location/Qualifiers	name	sequences	the	the	inst
		1. .167046					
		/organism="Mus musculus"					
		/mol_type="genomic DNA"					
		/db_xref="taxon:10090"					
		/chromosome="9"					
		/clone="RP24-511A23"					
		/clone_lib="RPCI-24"					
		16685_16712					
	misc_feature	/note="Sequence derived					
		31565_31779					
	unsure	/note="Unresolved simple					
		46721_46808					
	unsure						

```

/note="Unresolved simple sequence repeat."
142336..142347
/note="Sequence derived from one plasmid subclone."

ORIGIN
Alignment Scores:
Pred. No.: 78.7 Length: 167046
Score: 50.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-18 (1-9) x AC158516 (1-167046)

Qy 1 TyxMetAlaaspMetValalaTropLeu 9
|||||
Db 109920 TACATGGCTGACATGGTGGCTT 109894

RESULT 9
AF063939 2333 bp mRNA linear ROD 01-JAN-2000
LOCUS Rattus norvegicus 574 oncofetal antigen homolog (574) mRNA,
DEFINITION complete cds.
ACCESSION AF063939
VERSION AF063939.1 GI:6650211
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 2333)
AUTHORS Ninkina,N.N. and Buchman,V.L.
TITLE Structure and expression of the rat 574 gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2333)
AUTHORS Buchman,V.L.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1998) School of Biomedical Sciences, University
of St. Andrews, Bute Medical Buildings, St. Andrews, Fife KY16 9TS,
UK
FEATURES
Source Location/Qualifiers
1..2333
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="newborn"
1..2333
/gene="574"
1..363
/gene="574"
364..1644
/gene="574"
/codon_start=1
/product="574 oncofetal antigen homolog"
/protein_id="AAF21770.1"
/db_xref="GI:6650212"
/translation="MPGAGSRGFSAGDGRLRLARLARLVLLGWVSAPSSSLPSSSTS
PAAFLTAGSQPPPAERCPAAECSEAPATVRCVNRNLLEVPADLPYYRNLFLTGNQ
MTVLPAGAFARQPPPLADLAVNLGNHLKEVGAGAFEHLPLGLRLDLSHNPULTNLSP
TAGNSVSVTPSPLLLELNLHIVPEDORQSGFEGMVAFGMWAAALRSGLAGRLGL
HLELASNFVLYLPDRLLQPLSLKHLDRNNLSVLSLYTASFRNLTHLSLHEDNAL
KVLHNSLAEWQGLAHRVFLDNPPWCDYMSLVMSWLKETEVPDKARLTCAPPEK
MNRGLDLITSSDLQDCAATLPQSLQTSYVFLGLALIGALIFLLVLYLNKRGTKKWMH
NRDADRCHMEGVHYRYEINADPSLTNLSSSDV"
1645..2333
/gene="574"
2315..2320
/gene="574"

3'UTR
polyA_signal
ORIGIN

```

Alignment Scores:

Pred. No.: 3.16 Length: 2333
 Score: 47.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 94.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-18 (1-9) x AF063939 (1-2333)

QY 1 TyxMetAlaaspMetValAlaTrieu 9

Db 1282 TACATGGCTGACATGGTGTCTGGCTT 1308

RESULT 10

BC087011 2361 bp mRNA linear ROD 13-DEC-2004
 Rattus norvegicus trophoblast glycoprotein, mRNA (cDNA clone
 MGC:93332 IMAGE:7193411), complete cds.

ACCESSION

BC087011

VERSION

BC087011.1 GI:56268819

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Muridae; Muridae; Muridae; Muridae;

1 (bases 1 to 2361)

Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Schectel,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Buffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smalish,D.E.,

Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

JOURNAL

PUBMED

2 (bases 1 to 2361)

Director MGC Project.

Direct Submission

Submitted (02-DEC-2004) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Howard Jacobs

cDNA Library Preparation: Express Genomics

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-sngc.stanford.edu>Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAC Plate: 186 Row: 0 Column: 24

This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13929143.

FEATURES

source

1..2361
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="MGC:93332 IMAGE:7193411"
 /tissue_type="Heart, rat (Brown Norway)"
 /clone_lib="NIH MGC_234"
 /lab_host="DH10B"
 /note="Vector: pExpress1"
 1..2361

gene

/gene="Tpbg"
 /note="synonym: 574"
 /db_xref="GeneID:83684"
 364..1544
 /db_xref="RGD:621453"
 /gene="Tpbg"
 /codon_start=1
 /product="Tpbg protein"
 /protein_id="AAH87011.1"
 /db_xref="GI:56268820"
 /db_xref="GeneID:83684"
 /db_xref="RGD:621453"

CDS

/translation="MPGAGSRGSPAGDGRRLRLRLALVLLGWVSASAPSSSPSSSTS
 PAAFLAGSAQPPPAECPCACCEAARTVCVNRNLLEVPADLPYVNLFLTNQ
 MTLVPAGAFAROPPLADLAVNLNSGNHLKEVGAGAFELGLRLDLNPLTNLSAF
 TFAGSVSVSPLELLIHLVPEPORQNGSPGMYAFEGMYAALRSGIALRGL
 HHLEASNHFLYLPDLDLQSLKHLDRNLSVLTYSFNLTHLSLHLELDNL
 KVLNLTAEWGLAHVRVFLDNNPWCDCYNADVMVSLKETEVVDPDKARTCAFPDK
 MNRLGLDITSSLDLQDQATLPQSLQTSYVFLGIVLALIGALFLVLLYLNRRKTKKMMH
 NIRDACRDHMEGYHYRYEINADPRLTNLSSNDV"

ORIGIN

Alignment Scores:
 Pred. No.: 3.2 Length: 2361
 Score: 47.00 Matches: 8
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 88.9% Mismatches: 0
 Query Match: 94.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-18 (1-9) x BC087011 (1-2361)

QY 1 TyxMetAlaaspMetValAlaTrieu 9

Db 1282 TACATGGCTGACATGGTGTCTGGCTT 1308

RESULT 11

LOCUS

AC128294/c

DEFINITION

Rattus norvegicus clone CH230-176H20, WORKING DRAFT SEQUENCE.

ACCESSION

AC128294

VERSION

AC128294.3

KEYWORDS

HTG; HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Muridae; Muridae; Muridae; Muridae;

1 (bases 1 to 210237)

Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,

Anyalebechi,V., Ayodeji,A., Ayodeji,M., Baca,E., Baden,H.,

Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,

Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,

Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,

Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,

Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,

Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,

Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,

REFERENCE

AUTHORS

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregis, B., Geer, K., Gill, R., Grady, A., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhsawa, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarnpunagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shivartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, J., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

Direct Submission

JOURNAL

Unpublished

2 (bases 1 to 210237)

Worley, K.C.

AUTHORS

JOURNAL

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 210237)

Rat Genome Sequencing Consortium.

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23265004.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GZGV

Center clone name: CH230-176H20

----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 201781 bases at least Q40
 Consensus quality: 203921 bases at least Q30
 Consensus quality: 205310 bases at least Q20
 Estimated insert size: 205531; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

source

1. 210237: contig of 210237 bp in length.

Location/Qualifiers

1. 210237

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-176H20"

1. 1142

/note="wgs end extension"

clone_end:T7"

2177. 144799

/note="clone_boundary"

clone_end:T7"

end sequence: BH360464"

complement (206062. 206961)

/note="clone_boundary"

clone_end:Sp6"

site:

end sequence: BH360465"

208907. 210237

/note="wgs end extension"

clone_end:Sp6"

ORIGIN

Alignment Scores:

Pred. No.: 504 Length: 210237

Score: 47.00 Matches: 8

Percent Similarity: 100.0% Conservative: 1

Best Local Similarity: 88.9% Mismatches: 0

Query Match: 94.0% Indels: 0

DB: 12 Gaps: 0

US-10-774-176-18 (1-9) x AC128294 (1-210237)

QY

1 TyrMetAlaAspMetValalaTrpLeu 9

|||||

Db 110600 TACATGGCTGACATGCTGCTTGGCTT 110574

RESULT 12

AC106962/c

LOCUS

DEFINITION

Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4

unordered pieces.

AC106962

AC106962.5 GI:25139469

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 239076)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

AC106962 239076 bp DNA linear HTG 20-NOV-2002
 Rattus norvegicus clone CH230-87110, WORKING DRAFT SEQUENCE, 4
 unordered pieces.

AC106962.5 GI:25139469

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 239076)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Blawie, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, J., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensehuwa, L., Louissegh, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokenleh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Kelly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rockey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, R., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE 2 (bases 1 to 239076)

AUTHORS

TITLE

Submitted (14-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 239076)

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 20, 2002 this sequence version replaced gi:22857070. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: Gopi
Center clone name: CH230-87110
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 228642 bases at least Q40
Consensus quality: 232269 bases at least Q30
Consensus quality: 234041 bases at least Q20
Estimated insert size: 231522; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 234710: contig of 234710 bp in length
* 234711 234810: gap of unknown length
* 234811 235924: contig of 1114 bp in length
* 235925 236024: gap of unknown length
* 236025 237314: contig of 1290 bp in length
* 237315 237414: gap of unknown length
* 237415 239076: contig of 1662 bp in length.

FEATURES

source

1..239076
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-87110"
234711..234810
/estimated_length=unknown
235925..236024
/estimated_length=unknown
237315..237414
/estimated_length=unknown

ORIGIN

Alignment Scores:

Pred. No.:	583	Length:	239076
Score:	47.00	Matches:	8
Percent Similarity:	100.0%	Conservative:	1
Best Local Similarity:	88.9%	Mismatches:	0
Query Match:	94.0%	Indels:	0
DB:	12	Gaps:	0

US-10-774-176-18 (1-9) x AC106962 (1-239076)

QY 1 TyrMetAlaAspMetValAlaTrpLeu 9

DB 15811 TACATGGCTGACATGGTGTCTTGGCTT 15785

RESULT 13

AX149553

LOCUS

Sequence 14 from Patent WO0136486.

ACCESSION AX149553

VERSION AX149553.1 GI:14347991

KEYWORDS

synthetic construct

ORGANISM

other sequences; artificial sequences.

linear

DNA

PAT 08-JUN-2001

```
REFERENCE 1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
         Ellard,F.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
FEATURES Oxford Biomedica (UK) Limited (GB)
source Location/Qualifiers
1..1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="5T4"
ORIGIN
Alignment Scores:
Pred. No.: 4.59 Length: 1263
Score: 45.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-18 (1-9) x AX149553 (1-1263)
Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
Db 901 CACATGGCAGACATGGTGGCCTGGCTC 927
RESULT 14
AX467371 1263 bp DNA linear PAT 16-JUL-2002
LOCUS AX467371
DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS
SOURCE Canis sp.
ORGANISM Canis sp.
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
         Canis.
REFERENCE 1
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
FEATURES Oxford Biomedica (UK) Limited (GB)
source Location/Qualifiers
1..1263
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"
ORIGIN
Alignment Scores:
Pred. No.: 4.59 Length: 1263
Score: 45.00 Matches: 8
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 88.9% Mismatches: 0
Query Match: 90.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-18 (1-9) x AX467371 (1-1263)
Qy 1 TyrMetAlaAspMetValAlaTrpLeu 9
Db 901 CACATGGCAGACATGGTGGCCTGGCTC 927
RESULT 15
AX328003S21 75778 bp DNA linear BCT 03-MAR-2004
LOCUS AX328003S21/c
DEFINITION Symbiont bacterium of Paederus fuscipes putative isocitrate
         dehydrogenase kinase/phosphatase gene, partial cds; and putative
         dioxigenase, putative pederin biosynthesis mixed type I polyketide
         synthase/nonribosomal peptide synthetase gene cluster, hypothetical
protein, putative RNA methylase, hypothetical protein, and putative
aldehyde dehydrogenase genes, complete cds.
ACCESSION AY328023 AY059471
VERSION AY328023.1 GI:446662898
KEYWORDS
SEGMENT
SOURCE
ORGANISM
21 of 22
symbiont bacterium of Paederus fuscipes
symbiont bacterium of Paederus fuscipes
Bacteria.
1 (bases 1 to 75778)
Piel,J.
A polyketide synthase-peptide synthetase gene cluster from an
uncultured bacterial symbiont of Paederus beetles
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14002-14007 (2002)
12381784
2 (bases 1 to 75778)
Piel,J., Hofer,I. and Hui,D.
Evidence for a symbiosis island involved in horizontal acquisition
of pederin biosynthetic capabilities by the bacterial symbiont of
Paederus fuscipes beetles
Paederus fuscipes beetles
J. Bacteriol. 186 (5), 1280-1286 (2004)
14973122
3 (bases 1 to 75778)
Piel,J.
Direct Submission
Submitted (10-OCT-2001) Department of Bioorganics, Max Planck
Institute for Chemical Ecology, Winzerlaer Str. 10, Jena 07745,
Germany
4 (bases 1 to 75778)
Piel,J., Hofer,I. and Hui,D.
Direct Submission
Submitted (19-JUN-2003) Department of Bioorganic Chemistry, Max
Planck Institute for Chemical Ecology, Winzerlaer Str. 10, Jena
07745, Germany
On Mar 3, 2004 this sequence version replaced gi:23307838.
Location/Qualifiers
1..75778
/organism="symbiont bacterium of Paederus fuscipes"
/mol_type="genomic DNA"
/db_xref="taxon:176282"
complement(<1..520)
/note="similar to PA1376 from Pseudomonas aeruginosa"
/codon_start=1
/transl_table=11
/product="putative isocitrate dehydrogenase
kinase/phosphatase"
/protein_id="AA547548.1"
/db_xref="GI:44662930"
/translation="MAKYSADDIASSILRGFDNYRQFRRITDGRARFQAHWOEA
QRASIERINDYEEKVTETVDTLTNIPADLLIEIRWPTTIRSAVITLIDVRLDDELA
TWFNSIFCSLFSHDNISDDTMFVHTTRPSLRTHARAPYRVRPEGKLRQTLEQIFDD
YRFDVPHENRERD"
802..71987
/note="symbiosis island"
complement(802..939)
/note="putative transposase pseudogene"
1087..1415
/note="hypothetical protein pseudogene, similar to PA1689
from Pseudomonas aeruginosa"
1911..2654
/codon_start=1
/transl_table=11
/product="putative dioxigenase"
/protein_id="AA547549.1"
/db_xref="GI:44662931"
/translation="MSWIYDCAVSALRWFEEDRISGSATLEADOLFQNAHFVSNWAE
IREASRATNHLHKIPRFHELMRAQKISANDGRDWMFVARAYGVTFQPNLKQPTL
ASTLACSPDLSASFILGEGKFPFHRGFRGVLGYLVITMPKHDDGIPAAVLMD
GEHRLEEGDFVLWDDTFEHSVWVNESRIVLLLDIKRRNMPDLDSIVSNIYYIR
LSTLLSRSRRLAILSRLPDKADDSLDQS"
2728..3076
/note="putative resolvase pseudogene"
complement(3307..3804)
misc_feature
misc_feature
misc_feature
CDS
misc_feature
misc_feature
```

misc_feature	/note="hypothetical protein pseudogene, similar to PA52664 from <i>Pseudomonas aeruginosa</i> " 4168. .4508
misc_feature	/note="hypothetical protein pseudogene, similar to PA4500 from <i>Pseudomonas aeruginosa</i> " complement(4918. .5095)
gene	/note="putative transposase pseudogene" complement(6372. .7310)
CDS	/gene="peda" complement(6372. .7310) /gene="peda" /codon_start=1 /transl_table=11 /product="putative methyltransferase" /protein_id="AA047557.1" /db_xref="GI:44662939"
gene	/translation="MLKQRLPECISSHVDGIGYVPIDLQKIFDFGAAPSKDAGSEAS NINSEKERLASHINQHYDTFFSEGTSLLVDGSDYRNIGYMWDTTITTOHASEKSL QADLLDITPEKSGRILDAACMGSTARHLLSEYPADNIAWINTSEKQIEATRNRVPGCC HAQVFNADVLDSEFGGFDNLICIEAAHFHETROKFLSEARRILRPGRLVLSDVLFS SERLEQYPIFPFSAI>NHLDNTEYRRLKIDTGFQSOVEIDVSDVWGAAHFIVAVKRVHE AFYKGLDITVOLTEMLSNYYOLNISTIKHCLFICAQK" complement(7324. .8718)
CDS	/gene="pedb" complement(7324. .8718) /gene="pedb" /codon_start=1 /transl_table=11 /product="putative FMN-dependent oxidoreductase" /protein_id="AA047558.1" /db_xref="GI:44662940"
gene	/translation="MITAEKLSRLFKEDYQVRVAYVAGAMAKAIGSSDLVIAMGRAG VYGFVGGMTLEBIEITAIQRIQQGLNKGPEYGINVLANLADLEKNHLDLILRYGV RHIEAAPQMTIPALVITFRLKGLHLDASGKQQAANTLMKVSRPEIAQLFSLPVPASI VSLLTTEGRINEQAELIARIVASDVCEADSGGTHDMATVSVLPRIIRLRDITLQOQ QYRQSPVRVSGAGIGITPPSAASFMLGAEFIVTGSINQCTVQAGTSETVKKMLQSI DVQDTAPAGDMFELSGIKVLKKYFVFPVRANRLYDLWRNNGSLQLAPTVRKELQ DKYFKRSPDDIYQSTERYKRAKSAHTEKAERDPLKMWALIFRWYFVHTWRLALEGHH EORTDYQIHTGPAAGNRRWVKDTPLEDWKNRDVTKMADYLMDACAEYLNKRILASLTIT NRMSFTDHAASVI" complement(8894. .9898) /gene="pedc" complement(8894. .9898) /gene="pedc" /codon_start=1 /transl_table=11 /product="putative acyltransferase" /protein_id="AA047559.1" /db_xref="GI:44662941"
gene	/translation="MKDLQNTONHPVVMFSGQSGOYFQWGRQLYEQDETFFHAWKS LDDNRYDIGSLDILYDTGHERSLPDRILHTHPALFWQYVALAKSLARGLPAPD FLIGASGEFTAISAGDTHVENTLNLIKQARLFDYCNAGAMLLVIDHTDFTSTPT AFSDCELAGINFCHRVGSPRTGILQTKSLTKQNIACOLLPSVITAFHSSMDRVH BIFTKQPELQRCFLHTPVSICALPVEQLTFRFSTYMHVIRQPIAFHIAITPFQS SPNAVYLDLGPAGNMAATKYNLPSSITHYRILPTMTFPGRDLENIEIARLRLLELDQR "
CDS	10062. .11111 /gene="pedd" 10062. .11111 /gene="pedd" /codon_start=1 /transl_table=11 /product="putative acyltransferase" /protein_id="AA047563.1" /db_xref="GI:44662945"
gene	/translation="MGQLDFRFPNIIIEAANDILGYSIKTLCLEDQORQLRLQYTYQV QAPRGMAAIIIGISADSVAGILABQGLTRIDIANYNAPTOTIISGLEADIRDAQVPE ALYVGNALTRHQIQCGGGLPDFVAGHSLGEYNALASGVFSPEDGLRLVQKRGDMLS SCQMYVPLNTSGAFHSRYMQSARDEPAQFLEAFEPDRDPOIPVYANTAKPYGTGVV ATLADQUTQVRWLDSDNRFLDQDQVTEFLRGPDVLUSKLVSEIRSMSKPYVPSEFAE ENSQOLVDENRMTCPISGRVVRVGYDDILVTKSRALLFHGRRAAIYMWNYQGYFALSE VEPLIEQQLVKEVW" 11341. .12162
gene	

```

/gene="pedB"
11341..12162
/gene="pedB"
/codon_start=1
/transl_table=11
/product="putative methyltransferase"
/protein_id="AA947560.1"
/db_xref="GI:44662942"
/translations="MOTAIADVBEKATVYDSAGQVGPILFGGHEWGYWDEVTGBGN
FANAARLAQIMIAKAPIKAGQKIFDMGQGFSGSALKLAKAKGCFVDGTTISKEQQLS
ATIFRAEOLQOEVRPETHGSALNIPEDQSYDGMFEFESIFHMGHKAHLHEAARLVLP
GSTILLTDLPLPESTEARKEFWEHRIHNSFVSREDYPELLASAEFELIIDDITDNV
MPWLEPKLKAIEIHRHPQVEAIIIPNDTEKAIDMLVLFETYSMENLQFMIVMAKKL"
12407..38212
/gene="pedF"
12407..38212
/gene="pedF"
/codon_start=1
/transl_table=11
/product="mixed type I polyketide synthase/nonribosomal
peptide synthetase"
/protein_id="AA947564.1"
/db_xref="GI:44662946"
/translations="MSVNIHQQLKETEDALLNNSGVAVTLDVESDKLRKREAEDSPK
AIAIVELGYFPQASVDSEWRHLDQDATLIEIPSRFDWRKYVDPPTGERPGSSCSK
WGGFIIDPIRGDFPAPFNIGAEATLIDPPQRLLLMSAYQTLNDAGVASQALQSKTGTV
FVALQNEVQLLADAGIDPGMYAQTCLLANRISYFPDMWGTSEVVDQCPGAARLV
HRAVSALIRNGEIELALVGAANLLRPFPVLISESQLSSTSVHSGFAQAGAAQLRAE
GVCSSLPLPLTKALADGPTIYASIKHSVAFNFGOGGASIAAPNVDSHVDLIKSCYQRA
RVDPQVRYIEAGMGNSVLADIWEOAFNRALDIAQQRVSLPGNCILISTLKPMQG
HMESASALGALFKVILSLTRTHIKTAHTVOHPMDIYQGPCNIAELGVAMPQMEGL
RLAGIHCYGMGVNAHLLEVESVAGYYDSSELGTVSLSLEHLVILVLSAKTSESRLRMA
RRQGFQKQADAPALURDIATYLQVRDAPEHRLALVDSQQQLIEGLCYLEERQPS
RIGQFQVQASQVASESLPFTEDDLAAVACWVAGGAVLWVPVPGKPKRRVRLPAP
FYDRAVYVDSVAVESRANSPAKPSMLSGERSIGDYLRKLGELVQVPPVRIDPQGH
LYDLGVDSIVAMKLNLNLAFAFGIPVVRGDRLLQYSTVQALSRHLAQVLRDQVSEGE
DESPROLMASRCSLSEGQGLWVLQQLASRMTAYNIPLCVRIAQVLDITALEAFAM
LLEQYPLITSFVFDNGELFRCHVAAALPFQWQETNTLDQAVRMRLKCLAKQPFEL
EKGPLRHLVLSCEGHDIILLCVHHIYVDFGSLFPVFGGLIQTYYQLISOGQTLAKST
QGEYADAVFLVMEKQALASAGQRRHAYRKLQSLGELPVLISFTDNPRDAQGRFTDGT
YGFQLDVNLRSKTRNFKAQGNLMTLFLKLLHRYSGQSELIIGMPEQGRSEER
FELGVGVYVNLPIRSGYGVTFLEAFARDLQLSMADMDHAYVPFVPMVRLDGRAPA
EDLAPIVOFAVEYQVFSQDLRLFNQSVRESIGVTFLEFVQEGEYELALEVREGEL
DFALNLIKPTLYRMATIAMAEHLILIAEHAIDAPLSCRELTMLSERERHLLLHEW
NATTEPYSCCFHLPEKQARWMPQAI AAI FQEURSYAELDERSERLAIYDQCCGVG
PNRIYAVCLERSLDMLVLTIGARSGAANLPIDPNYPDRFLPMLSDSQALLLTCEG
LRDKTAATVYSQVAGERLQIVAMDGHWPTEQARTSELMQRDPRNLAVITYTSVSGTG
IPKGMIEHRSINFLVFLYMSNPLGRANDRLIAVTTYCFDIAGLELLVLLVCAGCCCI
CATDKLNDSEALQGEIERLQPTVMQATPSTWTLLPHFGGNNRQGVKILCGGBELPPLP

```

Search completed: May 27, 2006, 19:36:01
Job time : 3371.6 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-17

Perfect score: 52

Sequence: 1 HMADMTWL 9

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB spool/US10774176/runat 26052006 091441 24976/app query.fasta 1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=spt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02h
-USER=US10774176 @CGN 1 1 2389 @runat 26052006 091441 24976 -NCPU=6 -ICPU=3
-NO MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 8.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*
- 14: geneseqn2005s.*
- 15: geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	299	10 ACD93536	Human col
2	52	100.0	453	5 AAS87174	DNA encod
3	52	100.0	475	13 ADU11677	Solid tum

4	52	100.0	927	6 ABT07721	Breast ca
5	52	100.0	927	8 ABX76333	Lung canc
6	52	100.0	927	10 ADB80503	Ovarian c
7	52	100.0	927	11 ADR38723	Cancer/an
8	52	100.0	973	8 AAD56198	Human LRR
9	52	100.0	1156	6 AAV99349	Human NOV
10	52	100.0	1263	3 AAA27058	Human 5T4
11	52	100.0	1331	8 AAD56199	Human LRR
12	52	100.0	2020	10 ADJ56299	Human cDN
13	52	100.0	2053	8 ACC51052	Human bla
14	52	100.0	2053	8 AAX76332	Lung canc
15	52	100.0	2053	8 AAD56197	Human LRR
16	52	100.0	2053	8 AAD56200	Human LRR
17	52	100.0	2053	11 ADN38721	Cancer/an
18	52	100.0	2053	12 ADL06473	Human tum
19	52	100.0	2053	12 ADN03961	Antipseori
20	52	100.0	2053	13 ADR25444	Breast ca
21	52	100.0	2053	13 ACN38510	Tumour-as
22	52	100.0	2053	13 ADV35098	Human cDN
23	52	100.0	2053	14 AED17761	Fibrotic
24	52	100.0	2338	5 AAS87175	DNA encod
25	52	100.0	2359	4 AAK94253	Human ful
26	52	100.0	2359	12 ADL30831	Full leng
27	52	100.0	2361	4 AAK94254	Human ful
28	52	100.0	2361	12 ADL26162	Human cDN
29	52	100.0	2361	12 ADL30833	Full leng
30	47	90.4	1263	4 AAF89736	Nucleotid
31	47	90.4	1263	6 ABK87174	CDNA enco
32	43	82.7	1260	6 ABK87175	CDNA enco
33	43	82.7	1260	10 ADB97513	Feline 5T
34	43	82.7	1260	10 ADB97452	DNA encod
35	41	78.8	206	13 ADU13519	Solid tum
C 36	41	78.8	393	14 ADV73884	Human col
C 37	41	78.8	1026	6 ABK51946	3'-untran
C 38	41	78.8	1026	9 ACD12556	Human G-p
C 39	41	78.8	1281	3 AAA27059	Mouse 5T4
40	41	78.8	2557	12 ADI26160	Human cDN
41	41	78.8	2557	12 ADI26158	Human cDN
42	41	78.8	4118	4 AAI61358	Human pol
43	41	78.8	4118	10 ADC32089	Human nov
44	41	78.8	4370	13 ADR07185	Full leng
45	41	78.8	4387	4 AAI59574	Human pol

ALIGNMENTS

RESULT 1	
ACD93536	
ID ACD93536 standard; cDNA; 299 BP.	
AC	
ACD93536;	
XX	
DT	
DT 23-SEP-2003 (first entry)	
XX	
DE Human colon cancer cell expressed cDNA #1948.	
XX	
KW Open reading frame detection; genome sequencing; colon cancer;	
KW breast cancer; population genome analysis; genetic shift; cancer;	
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;	
KW agriculture; food crop genome; resistance gene; retrovirus;	
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;	
KW gene; ss.	
XX	
OS Homo sapiens.	
XX	
PN US2002155438-A1.	
XX	
PD 24-OCT-2002.	
XX	
PF 27-SEP-1999; 99US-00406117.	
XX	
PR 20-NOV-1998; 98US-00196716.	
XX	

PA (SIMP/) SIMPSON A J G.
 PA (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 XX
 XX Simpson AJG, Neto ED, Brentani RR;
 XX WPI; 2003-182626/18.
 XX
 XX Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 XX
 XX Example 9; Page 302; 959pp; English.
 XX
 CC The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (I) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. Further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a foetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 XX
 SQ Sequence 299 BP; 75 A; 84 C; 78 G; 62 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.389 Length: 299
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-17 (1-9) x ACD93536 (1-299)
 QY 1 HisMetAlaAspMetValThrTrpLeu 9
 Db 195 CACATGGCAGACATGGTGACCTGGCTC 221

RESULT 2
 AAS87174
 ID AAS87174 standard; cDNA; 453 BP.
 XX AAS87174;
 XX 13-FEB-2002 (first entry)
 XX

DE DNA encoding novel human diagnostic protein #22978.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; sa.
 XX
 OS Homo sapiens.
 XX
 FN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG22987.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 22978; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have application in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94584 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 453 BP; 108 A; 111 C; 113 G; 121 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.613 Length: 453
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 5 Gaps: 0

US-10-774-176-17 (1-9) x AAS87174 (1-453)
 QY 1 HisMetAlaAspMetValThrTrpLeu 9
 Db 88 CACATGGCAGACATGGTGACCTGGCTC 114

RESULT 3
 ADU11677
 ID ADU11677 standard; DNA; 475 BP.
 XX ADU11677;
 AC

XX 27-JAN-2005 (first entry)
 XX Solid tumour prognosis gene seqid 2116.
 XX cytostatic; gene therapy; expression profile; solid tumour;
 KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
 XX Unidentified.
 OS WO2004097052-A2.
 XX 11-NOV-2004.
 PD 29-APR-2004; 2004WO-US013587.
 PF 29-APR-2003; 2003US-0466067P.
 PR 23-JAN-2004; 2004US-0538246P.
 XX (AMHP) WYETH.
 PA (STRA/) STRAHS A.
 XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
 PI Immerman F, Dörner AJ;
 XX WPI; 2004-804779/79.
 DR A method, useful for prognosing and treating solid tumor, comprises
 PT comparing an expression profile of a gene expressed in peripheral blood
 PT mononuclear cells to a reference expression profile of a gene.
 XX Disclosure; Page; 11pp; English.
 XX The invention describes a method comprising comparing an expression
 CC profile of at least one gene in a peripheral blood sample of a patient to
 CC at least one reference expression profile of the at least one gene, where
 CC the patient has a solid tumour, and each of the gene is differentially
 CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
 CC of patients as compared to PBMCs of a second class of patients, where
 CC each of the first and second classes is a subcluster formed by an
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of
 CC a population of patients who have the solid tumour, and where the
 CC majority of the first class of patients has a first clinical outcome, and
 CC the majority of the second class of patients has a second clinical
 CC outcome. Also described are: a system comprising (i) a memory or a
 CC storage medium including data that represent an expression profile of at
 CC least one gene in a peripheral blood sample of a patient who has a solid
 CC tumour, (ii) at least another storage medium including data that
 CC represent at least one reference expression profile of the gene, (iii) a
 CC program capable of comparing the expression profile to the reference
 CC expression profile, and (iv) a processor capable of executing the
 CC program, where expression levels of the gene in peripheral blood
 CC mononuclear cells of patients who have the solid tumour correlate with
 CC clinical outcomes of the patients; and a nucleic acid or protein array
 CC comprising concentrated probes for solid tumour prognosis genes, where
 CC each of the solid tumour prognosis genes is differentially expressed in
 CC PBMCs of a first class of patients as compared to PBMCs of a second class
 CC of patients, where both the first and second classes of patients have a
 CC solid tumour, and where the first class of patients has a first clinical
 CC outcome, and the second class of patients has a second clinical outcome.
 CC The method, system, and array are useful for prognosing and treating
 CC solid tumours. This sequence represents a solid tumour prognosis gene of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0.646 Length: 475
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 13 Gaps: 0
 US-10-774-176-17 (1-9) x ADU11677 (1-475)
 QY 1 HisMetAlaAspMetValThrTrpLeu 9
 DB 180 CACATGGCAGACATGGTGACCTGGCTC 206
 RESULT 4
 ABT07721
 ID ABT07721 standard; DNA; 927 BP.
 XX AC ABT07721;
 XX 14-NOV-2002 (first entry)
 DT DE Breast cancer-associated gene sequence 29.
 XX Gene; ds; breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.
 XX Unidentified.
 OS WO200259377-A2.
 XX 01-AUG-2002.
 PD 24-JAN-2002; 2002WO-US002242.
 PF 24-JAN-2001; 2001US-0263965P.
 PR 02-FEB-2001; 2001US-0265928P.
 PR 09-APR-2001; 2001US-00829472.
 PR 09-APR-2001; 2001US-0282698P.
 PR 04-MAY-2001; 2001US-0288590P.
 PR 29-MAY-2001; 2001US-0294443P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA Mack DH, Gish KC, Afar D;
 XX WPI; 2002-583738/62.
 DR N-PSDB; AB005564.
 XX Detecting a breast cancer-associated transcript in a patient's cell,
 PT useful for diagnosing breast cancer, comprises contacting a biological
 PT sample with a polynucleotide that selectively hybridizes with breast
 PT cancer nucleic acids.
 XX Claim 9; Page 372; 414pp; English.
 PS The invention comprises a method of detecting a breast cancer-associated
 CC transcript in a cell from a patient. The method of the invention involves
 CC contacting a biological sample from the patient with a nucleotide that
 CC hybridises to one of the 69 breast cancer-associated gene sequences shown
 CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences AB07693 - AB07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.34 Length: 927
 Score: 52.00 Matches: 9

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-17 (1-9) x ABT07721 (1-927)

QY 1 HisMetAlaAspMetValThrTrpLeu 9

Db 559 CACATGGCAGACATGGTGACCTGGCTC 585

RESULT 5

ABX76333

ID ABX76333 standard; DNA; 927 BP.

XX AC ABX76333;

DT 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #197.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334370P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX P-PSDB; ABUS5604.

PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX Claim 22; Page 336; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences

CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 SQ

Alignment Scores:

Pred. No.: 1.34 Length: 927
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-17 (1-9) x ABX76333 (1-927)

QY 1 HisMetAlaAspMetValThrTrpLeu 9

Db 559 CACATGGCAGACATGGTGACCTGGCTC 585

RESULT 6

ADB80503

ID ADB80503 standard; DNA; 927 BP.

XX AC ADB80503;

XX 04-DEC-2003 (first entry)

DE Ovarian cancer-associated transcript #34.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;

KW post-operative chemotherapy; radiation therapy; tumour prognosis;

KW pre-cancerous lesion detection; ds; gene.

XX Homo sapiens.

XX Key

Location/Qualifiers

CDS 1..927

/*tag= a

XX WO2002102235-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019297.

XX 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Gish KC;

XX WPI; 2003-167431/16.

XX P-PSDB; ADB80504.

XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

XX Claim 10; Page 297; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,

CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.

XX Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.34 Length: 927
Score: 52.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-17 (1-9) x ADB80503 (1-927)

QY 1 HisMetAlaAspMetValThrTrpLeu 9
|||||
Db 559 CACATGGCAGACATGGTGACCTGGCTC 585

RESULT 7

ADN38723
ID ADN38723 standard; cDNA; 927 BP.

XX AC ADN38723;

XX DT 17-JUN-2004 (first entry)

XX DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.

XX OS Homo sapiens.

XX PN WO2003042661-A2.

XX PD 22-MAY-2003.

XX PF 13-NOV-2002; 2002WO-US036810.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 21-NOV-2001; 2001US-0332464P.

XX PR 29-NOV-2001; 2001US-0334393P.

XX PR 03-DEC-2001; 2001US-0335394P.

XX PR 14-DEC-2001; 2001US-0340376P.

XX PR 08-JAN-2002; 2002US-0347211P.

XX PR 10-JAN-2002; 2002US-0347349P.

XX PR 08-FEB-2002; 2002US-0355250P.

XX PR 13-FEB-2002; 2002US-0356714P.

XX PR 20-FEB-2002; 2002US-0359077P.

XX PR 29-MAR-2002; 2002US-036809P.

XX PR 04-APR-2002; 2002US-0370110P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PR 05-JUN-2002; 2002US-0386614P.

XX PR 16-JUL-2002; 2002US-0396839P.

XX PR 22-JUL-2002; 2002US-039775P.

DR P-PSDB; ADN38724.

XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

XX PS Claim 8; SEQ ID NO 41; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.

XX SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.34 Length: 927
Score: 52.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 11 Gaps: 0

US-10-774-176-17 (1-9) x ADN38723 (1-927)

QY 1 HisMetAlaAspMetValThrTrpLeu 9
|||||
Db 559 CACATGGCAGACATGGTGACCTGGCTC 585

RESULT 8

AAD56198

ID AAD56198 standard; DNA; 973 BP.

XX AC AAD56198;

XX DT 07-AUG-2003 (first entry)

XX DE Human LRRCAPS related DNA #5.

XX Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.

XX OS Homo sapiens.

XX PN WO2003035831-A2.

XX PD 01-MAY-2003.

XX PF 21-OCT-2002; 2002WO-US033540.

XX PR 22-OCT-2001; 2001US-0338733P.

XX PR 15-FEB-2002; 2002US-0357600P.

XX PR 01-MAR-2002; 2002US-0361196P.

XX PA (EXEL-) EXELIXIS INC.

XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;

XX PI Francis-Lang H, Friedman L;

XX DR WPI; 2003-421410/39.

XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.
 XX
 PS Example 5; Page 74-75; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a
 CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS related DNA

XX SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.42 Length: 973
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-17 (1-9) x AAD56198 (1-973)

Qy 1 HisMetAlaApMetValThrTrpLeu 9

Db 574 CACATGGCAGACATGGTGACCTGGCTC 600

RESULT 9

ABV99349

ID ABV99349 standard; DNA; 1156 BP.

XX AC ABV99349;

XX DT 27-JAN-2003 (first entry)

XX DE Human NOV8a coding sequence.

XX Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiant; hemostatic; neuroprotective; anorectic;
 KW neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antiinfertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

XX OS Homo sapiens.

XX PN WO200272771-A2.

XX PD 19-SEP-2002.

XX PF 08-MAR-2002; 2002WO-US007288.

XX PR 08-MAR-2001; 2001US-0274101P.

PR 08-MAR-2001; 2001US-0274194P.

PR 08-MAR-2001; 2001US-0274281P.

PR 08-MAR-2001; 2001US-0274332P.

PR 08-MAR-2001; 2001US-0274849P.

PR 12-MAR-2001; 2001US-0275235P.

PR 13-MAR-2001; 2001US-0275578P.

PR 13-MAR-2001; 2001US-0275579P.

PR 14-MAR-2001; 2001US-0275601P.

PR 14-MAR-2001; 2001US-0276000P.

PR 16-MAR-2001; 2001US-0276776P.
 PR 19-MAR-2001; 2001US-0276994P.
 PR 20-MAR-2001; 2001US-0277239P.
 PR 20-MAR-2001; 2001US-0277321P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277751P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0278999P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0279959P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 02-APR-2001; 2001US-0280900P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 02-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 03-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 30-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 21-NOV-2001; 2001US-0333272P.
 PR 03-DEC-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 04-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.

(CURA-) CURAGEN CORP.

Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 Boldog FL, Li L, Zerkhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 Zhong M;

WPI; 2002-732824/79.

P-PSDB; ABP70071.

New NOVX polypeptides and polynucleotides, useful for preventing,
 diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 disorders, and asthma.

Claim 16; Page 114-115; 619pp; English.

CC The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, haematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods

XX
 SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.71 Length: 1156
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-17 (1-9) x ABV99349 (1-1156)

Qy 1 HisMetAlaAspMetValThrTrpLeu 9
 Db 790 CACATGGCAGACATGGTGACCTGGCTC 816

RESULT 10

AA27058
 ID AAA27058 standard; DNA; 1263 BP.
 XX
 AC AAA27058;
 XX
 XX 22-AUG-2000 (first entry)
 XX Human 5T4 tumour-associated antigen gene.
 XX
 XX Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
 KW ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200029428-A2.
 XX
 XX 25-MAY-2000.
 XX
 XX 18-NOV-1999; 99WO-GB003859.
 XX
 XX 18-NOV-1998; 98GB-00025303.
 PR 27-JAN-1999; 99GB-00001739.
 PR 30-JUL-1999; 99GB-00017995.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX

PI Carroll MW, Myers KA;

DR WPI; 2000-387735/33.

XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 PT response useful in vaccinating against and in treating tumors.

XX Example 2; Page 78; 79pp; English.

CC The present sequence encodes the human 5T4 tumour-associated antigen
 CC (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in
 CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC present sequence. The 5T4 antigen was shown to be effective at eliciting
 CC an immunotherapeutic anti-tumour response. Both the nucleic acid encoding
 CC the antigen and the antigen itself can be used to elicit an immune
 CC response, preferably CTL or an antibody response in a subject
 XX
 SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.89 Length: 1263
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0

US-10-774-176-17 (1-9) x AAA27058 (1-1263)

Qy 1 HisMetAlaAspMetValThrTrpLeu 9
 Db 901 CACATGGCAGACATGGTGACCTGGCTC 927

RESULT 11

AA27058
 ID AAD56199 standard; DNA; 1331 BP.
 XX
 AC AAD56199;
 XX
 XX 07-AUG-2003 (first entry)
 XX Human LRRCAPS related DNA #6.
 DE
 XX Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.
 XX
 XX Homo sapiens.
 XX
 XX WO2003035831-A2.
 XX
 XX 01-MAY-2003.
 XX
 XX 21-OCT-2002; 2002WO-US033540.
 XX
 XX 22-OCT-2001; 2001US-0338733P.
 PR 15-FEB-2002; 2002US-0357600P.
 PR 01-MAR-2002; 2002US-0361196P.
 XX
 XX (EXEL-) EXELIXIS INC.
 XX
 XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;
 XX
 XX WPI; 2003-421410/39.

XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.

XX Disclosure; Page 75-76; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified Leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful

CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
SQ Sequence 1331 BP; 252 A; 447 C; 369 G; 263 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2 Length: 1331
Score: 52.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-17 (1-9) x AAD56199 (1-1331)

Qy 1 HisMetAlaAspMetValThrTripleu 9
|||
Db 931 CACATGGCAGACATGGTGACCTGGCTC 957

RESULT 12

ADJ56299
ID ADJ56299 standard; cDNA; 2020 BP.

XX
AC ADJ56299;

XX 06-MAY-2004 (first entry)

XX Human cDNA differentially expressed in MYCN activated cells SeqID 105.

XX human; differential expression; transactivator; proto-oncogene;

KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;

KW MYCN activated cell.

XX Homo sapiens.

XX US2003119009-A1.

XX 26-JUN-2003.

XX 25-FEB-2002; 2002US-00084817.

XX 23-FEB-2001; 2001US-0270784P.

XX (STUA/) STUART S G.

XX (NUCH/) NUCHTERN J G.

XX (PLON/) PLON S E.

XX (SHOH/) SHOHET J M.

XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;

XX WPI; 2003-635698/60.

XX New genes regulated by MYCN activation, useful in gene therapy,
XX particularly for treating a subject with e.g. neuroblastoma or other
XX cancers, or for diagnosing, staging or monitoring the treatment of the
XX cancer.

XX Claim 1; SEQ ID NO 105; 27pp; English.

XX This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
XX polynucleotide sequences that exhibit differential expression patterns in
XX cells activated by the transactivator MYCN, where MYCN is a proto-
XX oncogene that is amplified in neuroblastoma cells and is common in small
XX cell lung cancer. The present invention describes these cDNA molecules
XX as useful for in hybridisation assays to detect expression of nucleic
XX acids (or complementary nucleic acids) in a present in a given sample, as
XX well as for screening assays by identifying molecules or compounds that
XX specifically bind the cDNA as a ligand and modulate function or activity.
XX Accordingly, these compositions exhibit cytostatic activity and can also
XX be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX that is differentially expressed in MYCN activated cells, given in an

CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.

XX Sequence 2020 BP; 451 A; 591 C; 493 G; 485 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.16 Length: 2020
Score: 52.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-17 (1-9) x ADJ56299 (1-2020)

Qy 1 HisMetAlaAspMetValThrTripleu 9
|||
Db 971 CACATGGCAGACATGGTGACCTGGCTC 997

RESULT 13

ACC51052

ID ACC51052 standard; cDNA; 2053 BP.

XX
AC ACC51052;

XX 12-JUN-2003 (first entry)

XX Human bladder cancer associated cDNA sequence SEQ ID NO:192.

XX Human; bladder cancer; cytostatic; gene therapy; vaccine; gene; ss.

XX Homo sapiens.

XX WO2003003906-A2.

XX 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US021338.

XX 03-JUL-2001; 2001US-0302814P.

XX 03-AUG-2001; 2001US-0310099P.

XX 08-NOV-2001; 2001US-0343705P.

XX 13-NOV-2001; 2001US-0350666P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Aziz N;

XX WPI; 2003-201532/19.

XX P-PSDB; ABR48236.

XX Detecting a bladder cancer-associated transcript in a cell from a
XX patient, comprises contacting a biological sample from the patient with a
XX bladder cancer-associated polynucleotide or antibody.

XX Claim 6; Page 296; 307pp; English.

XX The present invention describes a method for detecting a bladder cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with a polynucleotide
XX that selectively hybridises to a sequence that is 80 % identical to a
XX table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
XX encode the human bladder cancer-associated proteins given in ABR48146 to
XX ABR48242). Bladder cancer-associated sequences from the present invention
XX have cytostatic activities, and can be used in antisense gene therapy and
XX in vaccine production. The method can be used for detecting a bladder
XX cancer-associated transcript in a cell from a patient. The method is
XX useful in diagnosing or treating bladder cancer and in screening for
XX compounds that modulate bladder cancer, such as hormones or antibodies.
XX The nucleic acid molecules from the present invention may be used in

CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications
 XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.22 Length: 2053
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-17 (1-9) x ACC51052 (1-2053)

Oy 1 HisMetAlaAspMetValThrTrpLeu 9

Db 985 CACATGGCAGACATGGTGACCTGGCTC 1011

RESULT 14

ABX76332

ID ABX76332 standard; DNA; 2053 BP.

AC ABX76332;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polynucleotide #196.

XX Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

XX 10-MAY-2001; 2001US-0290492P.

XX 09-NOV-2001; 2001US-0339245P.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0334370P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX P-PSDB; ABU56603.

PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX Claim 22; Page 335; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by

CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention

XX Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.22 Length: 2053
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-17 (1-9) x ABX76332 (1-2053)

Oy 1 HisMetAlaAspMetValThrTrpLeu 9

Db 985 CACATGGCAGACATGGTGACCTGGCTC 1011

RESULT 15

AAD56197

ID AAD56197 standard; DNA; 2053 BP.

XX AAD56197;

XX 07-AUG-2003 (first entry)

XX Human LRRCAPS DNA #11.

XX Human; p53 pathway; Leucine rich repeat capricious related protein;
 KW LRRCAPS; cancer; gene therapy; ds.

XX Homo sapiens.

XX WO2003035831-A2.

XX 01-MAY-2003.

XX 21-OCT-2002; 2002WO-US033540.

XX 22-OCT-2001; 2001US-0338733P.

XX 15-FEB-2002; 2002US-0357600P.

XX 01-MAR-2002; 2002US-0361196P.

XX (EXEL-) EXELIXIS INC.

XX Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
 PI Francis-Lang H, Friedman L;

XX WPI; 2003-421410/39.

XX Identifying a candidate p53 pathway-modulating agent for treating cancer
 PT comprises contacting an assay system comprising a purified leucine rich
 PT repeat, capricious related polypeptide or nucleic acid with a test agent.

XX Example 5; Page 73-74; 99pp; English.

XX The invention relates to a method of identifying a candidate p53 pathway
 CC modulating agent. The method involves contacting an assay system
 CC comprising a purified Leucine rich repeat, capricious related (LRRCAPS)
 CC polypeptide or nucleic acid or its fragment with a test agent and
 CC detecting a test agent-biased activity, where a difference between the
 CC test agent-biased activity and the reference activity identifies the test
 CC agent as a candidate p53 pathway modulating agent. The method is useful
 CC for identifying a candidate p53 pathway-modulating agent for preparing a

CC composition for diagnosing or treating cancer. The invention is useful in
 CC gene therapy. The present sequence is human LRRCAPS DNA

XX
 SQ Sequence 2053 BP; 461 A; 602 C; 499 G; 491 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.22 Length: 2053
 Score: 52.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-17 (1-9) x AAD56197 (1-2053)

Qy 1 HisMetalAspMetValThrTyrLeu 9
 |||||
 Db 985 CACATGGCAGACATGGTGCCTGGCTC 1011

Search completed: May 27, 2006, 10:37:47
 Job time : 381.5 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds
(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-17
Perfect score: 52
Sequence: 1 HMA DMVTWL 9

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abs/ABSSWEB pool/US10774176/runat_26052006.091443_24987/app query.fasta_1
-DB=GenEmbl -QWTF=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=pto -NORM=ext -HRAPISE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04
-USER=US10774176 @CGN 1.1 7534 @runat_26052006.091443_24987 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_scs:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_btg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Match	Length	Description

c 1	52	100.0	421	7 HSPA32B9
2	52	100.0	475	2 CQ920916
3	52	100.0	927	2 AX829164

4	52	100.0	1156	2	DD161112
5	52	100.0	1263	2	BD249731
6	52	100.0	1263	2	AX025011
7	52	100.0	1283	2	AX316086
8	52	100.0	2053	2	CQ731678
9	52	100.0	2053	2	DD174290
10	52	100.0	2053	5	HS5740A
11	52	100.0	2359	2	BD127282
12	52	100.0	2359	2	CQ782724
13	52	100.0	2359	5	AK074786
14	52	100.0	2361	2	AX961916
15	52	100.0	2361	2	BD127283
16	52	100.0	2361	2	CQ782726
17	52	100.0	2361	5	AK074790
18	52	100.0	2379	5	BC037161
19	52	100.0	2714	5	AB168308
20	52	100.0	5551	5	HSA012159
21	52	100.0	121909	5	HSJ492P14
22	47	90.4	1263	2	AX149553
23	47	90.4	1263	2	AX467371
24	46	88.5	157265	12	AC149011
25	46	88.5	177646	12	AC149257
26	45	86.5	137431	14	CR974572
27	45	86.5	216079	12	CT573045
28	44	84.6	235040	12	AC111883
29	43	82.7	1260	2	AX467373
30	43	82.7	1260	2	AX821533
31	43	82.7	1260	2	AX821548
32	43	82.7	163469	5	AC080009
33	43	82.7	190900	12	AC134803
34	43	82.7	209820	6	AC138109
35	43	82.7	216812	12	AC172223
36	43	82.7	218573	6	AC115358
37	43	82.7	222444	12	AC097876
38	43	82.7	223497	12	AC115162
39	43	82.7	243801	12	AC172361
40	42	80.8	2333	6	AF063939
41	42	80.8	2361	6	BC087011
42	42	80.8	32799	4	SPBC1703
43	42	80.8	42314	13	AC158495
44	42	80.8	91949	12	AC159552
45	42	80.8	91949	12	AC159552

ALIGNMENTS

RESULT 1	HSPA32B9/c	421 bp	DNA	linear	STS 21-MAY-1998
LOCUS	H.sapiens flow-sorted chromosome 6	HindIII	fragment	SC6pA32B9,	
DEFINITION	sequence tagged site.				
ACCESSION	294208	1	GI:1945202		
VERSION	294208.1				
KEYWORDS	STS; single read.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mungall,A.J., Huckle,E., Langford,C., Ross,M.T. and Hunt,S.E.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-APR-1997) The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact: humquery@sanger.ac.uk				
COMMENT	Vector: pBSISK+.				
FEATURES	Location/Qualifiers				
source	1..421				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="6"				
	/clone="SC6pA32B9"				

```

/sex="female"
/tissue type="EBV lymphoblastoid cell line"
/clone_lib="SC6pa"
/dev stages="adult"
/note="The estimated purity of the flow-sorted chromosome
6 library is >97%"

ORIGIN

Alignment Scores:
Pred. No.:      0.012      Length:      421
Score:          52.00      Matches:    9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:     100.0%      Indels:    0
DB:              7           Gaps:      0

US-10-774-176-17 (1-9) x HSPA32B9 (1-421)

Qy      1 HisMetAlaAapMetValThrTripleu 9
|||||
Db      412 CACATGGCAGACATGGTGACNTGGCTC 386

RESULT 2
LOCUS   CQ920916                475 bp      DNA
DEFINITION Sequence 2116 from Patent WO2004097052.
ACCESSION CQ920916
VERSION   CQ920916.1 GI:56210857
KEYWORDS Homo sapiens (human)
SOURCE   Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
AUTHORS Strahs,A., Immerman,F. and Dörner,A.J.
TITLE Methods for prognosis and treatment of solid tumors
JOURNAL Patent: WO 2004097052-A 2116 11-NOV-2004;
Wyeth (US); Burczynski, Michael E. (US)
FEATURES             Location/Qualifiers
source               1..475
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:      0.0138     Length:      475
Score:          52.00      Matches:    9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:     100.0%      Indels:    0
DB:              2         Gaps:      0

US-10-774-176-17 (1-9) x CQ920916 (1-475)

Qy      1 HisMetAlaAapMetValThrTripleu 9
|||||
Db      180 CACATGGCAGACATGGTGACCTGGCTC 206

RESULT 3
LOCUS   AX829164                927 bp      DNA
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION   AX829164.1 GI:39838931
KEYWORDS Homo sapiens (human)
SOURCE   Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

AUTHORS Carroll, M.W. and Myers, K.A.
 TITLE Polypeptide
 JOURNAL OXFORD BIOMEDICA LTD
 COMMENT OS Homo sapiens (human)
 PN JP 2002530060-A/1
 PD 17-SEP-2002
 PF 18-NOV-1999 JP 2000582415
 PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
 30-JUL-1999 GB 9917995.4
 PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
 PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K7/06, C07K14/065,
 PC C07K19/00,
 CC C12N15/00,
 CC Polypeptide
 FH Key
 FT source 1. .1263 Location/Qualifiers
 FT /organism='Homo sapiens (human)'.
 FEATURES
 source Location/Qualifiers
 1. .1263
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: Length: 1263
 Pred. No.: 0.0451 Matches: 9
 Score: 52.00
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0
 US-10-774-176-17 (1-9) x BD249731 (1-1263)
 QY 1 HisMetAlaAspMetValThrTrpLeu 9
 DB 901 CACATGGCAGACATGGTACCTGGCTC 927
 RESULT 6
 AX025011 1263 bp DNA linear PAT 15-SEP-2000
 LOCUS
 DEFINITION Sequence 1 from Patent WO0029428.
 ACCESSION AX025011
 VERSION AX025011.1 GI:10184932
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1
 AUTHORS Carroll, M.W. and Myers, K.A.
 TITLE Polypeptide
 JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
 CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
 BIOMEDICA LTD (GB)
 FEATURES
 source Location/Qualifiers
 1. .1263
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores: Length: 1263
 Pred. No.: 0.0451 Matches: 9
 Score: 52.00
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0
 US-10-774-176-17 (1-9) x AX025011 (1-1263)

```

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      901 CACATGGCAGACATGGTGACCTGGCTC 927

RESULT 7
LOCUS   AX316086               1263 bp      DNA      linear      PAT 14-DEC-2001
DEFINITION
Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION   AX316086.1 GI:17899278
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.

REFERENCE
AUTHORS  Carroll,M.W. and Myers,K.A.
TITLE    5t4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL  Patent: Ep 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)

FEATURES
source   1..1263
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      0.0451      Length:      1263
Score:          52.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             2          Gaps:         0

US-10-774-176-17 (1-9) x AX316086 (1-1263)

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      901 CACATGGCAGACATGGTGACCTGGCTC 927

RESULT 8
LOCUS   CQ731678               2053 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION
Sequence 17612 from Patent WO02068579.
ACCESSION CQ731678
VERSION   CQ731678.1 GI:42308932
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.

REFERENCE
AUTHORS  Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE    Kits, such as nucleic acid arrays, comprising a majority of
          humanexons or transcripts, for detecting expression and other uses
          thereof
JOURNAL  Patent: WO 02068579-A 17612 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
source   1..2053
          Location/Qualifiers
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      0.0812      Length:      2053
Score:          52.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             2          Gaps:         0

US-10-774-176-17 (1-9) x CQ731678 (1-2053)

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      985 CACATGGCAGACATGGTGACCTGGCTC 1011

RESULT 10
LOCUS   HS5T4OA               2053 bp      RNA      linear      PRI 18-APR-2005
DEFINITION
Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.
ACCESSION Z29083
VERSION   Z29083.1 GI:435654
KEYWORDS  5T4 gene; 5T4 oncofoetal antigen.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.

REFERENCE
1 (bases 1 to 2053)

```

```

Best Local Similarity: 100.0%      Mismatches: 0
Query Match:          100.0%      Indels:      0
DB:                   2          Gaps:         0

US-10-774-176-17 (1-9) x CQ731678 (1-2053)

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      985 CACATGGCAGACATGGTGACCTGGCTC 1011

RESULT 9
LOCUS   DD174290               2053 bp      DNA      linear      PAT 19-DEC-2005
DEFINITION
METHODS OF DIAGNOSIS OF BLADDER CANCER, COMPOSITIONS AND METHODS OF
SCREENING FOR MODULATORS OF BLADDER CANCER.
ACCESSION DD174290
VERSION   DD174290.1 GI:83972266
KEYWORDS  JP 2005514908-A/102.
SOURCE   Homo sapiens
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.

REFERENCE
1 (bases 1 to 2053)
AUTHORS  Mack,D.H. and Agiz,N.
TITLE    METHODS OF DIAGNOSIS OF BLADDER CANCER, COMPOSITIONS AND METHODS OF
          SCREENING FOR MODULATORS OF BLADDER CANCER
JOURNAL  Patent: JP 2005514908-A 102 26-MAY-2005;
Protein Design Labs Inc

COMMENT   OS Homo Sapiens
          PN JP 2005514908-A/102
          PD 26-MAY-2005
          PF 03-JUL-2002 JP 2003509925
          PR 12-APR-2002 US 60/372246,13-NOV-2001 US 60/350666, PR
          03-AUG-2001 US 60/310099,08-NOV-2001 US 60/343705, PR
          03-JUL-2001 US 60/302814
          PI david h mack,natasha agiz
          CC

FEATURES
source   1..2053
          Location/Qualifiers
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      0.0812      Length:      2053
Score:          52.00      Matches:      9
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             2          Gaps:         0

US-10-774-176-17 (1-9) x DD174290 (1-2053)

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      985 CACATGGCAGACATGGTGACCTGGCTC 1011

RESULT 10
LOCUS   HS5T4OA               2053 bp      RNA      linear      PRI 18-APR-2005
DEFINITION
Homo sapiens 5T4 gene for 5T4 oncofoetal antigen.
ACCESSION Z29083
VERSION   Z29083.1 GI:435654
KEYWORDS  5T4 gene; 5T4 oncofoetal antigen.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homiidae; Homo.

REFERENCE
1 (bases 1 to 2053)

```

AUTHORS Myers,K.A., Rahi-Saund,V., Davison,M.D., Young,J.A., Cheater,A.J.
and Stern,P.L.
TITLE Isolation of a cDNA encoding 574 oncofetal trophoblast
glycoprotein. An antigen associated with metastasis contains
leucine-rich repeats
J. Biol. Chem. 269 (12), 9319-9324 (1994)
JOURNAL 8132670
PUBMED
REFERENCE 2 (bases 1 to 2053)
AUTHORS Myers,K.A.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1993) Myers K. A., Paterson Institute for Cancer
Research, Immunology, Wilmslow Road, Manchester, M20 9BX, UK
FEATURES
Location/Qualifiers
source
1..2053
/organism="Homo sapiens"
/mol_type="other RNA"
/db_xref="taxon:9606"
/sex="female"
/tissue_type="placenta"
/clone_lib="lambda gt11 library of J. Milan"
62..372
/product="LRR N-terminal flank"
/label=N-flank
85..1347
/experiment="experimental evidence, no additional details
recorded"
/codon_start=1
/product="574 oncofoetal antigen"
/protein_id="CAA82324.1"
/db_xref="GI:435655"
/db_xref="GOA:Q13641"
/db_xref="InterPro:IPR000372"
/db_xref="InterPro:IPR000483"
/db_xref="InterPro:IPR01611"
/db_xref="UniProtKB/TrEMBL:Q13641"
/translation="MPGCSRGPAAGDGRRLRLRLALVLLGVSSSSPTSSASSPSS
APFLASVAPQPLPDPQPCALCESEAAATKCVNRNLTEVPTDLPATVRNLFUGNQ
LAVLPAGAFARPPPLAEALNLSGSLRDEVRAGAFELPSLRQLDLSHNLADLSPP
AFSGNSASVAPSLVLIILNHVPPEDQRNQRFEGMVVAALLAGRALQGLRLLELA
SNHFLYLPRDLVLAQLPSLRHLDLNNLSVLSVTYSFRNLTHLESHLSDNALKVLHNG
TLAEQLGLPHTRVFLDNNPWCDCMADMVTLKETEVVQCKDRLTCAYPEKMRNVL
LELSALDCDPIPLPSSLOTSYVFLGIVLALIGAFLVLVLYLNKGIKKMHNIRDAC
RDHMEGHYHYREINADPRLTNLSSNDV"
130..171
373..966
/product="Leucine rich repeat region"
/label=LRRs
966..1119
/product="LRR C-terminal flank"
/label=C-flank
1153..1215
/product="transmembrane peptide"
/standard_name="transmembrane region"
/function="Anchorage of the protein to the cell membrane"

sig_peptide
misc_RNA
misc_RNA
misc_RNA
misc_RNA

ORIGIN
Alignment Scores:
Pred. No.: 0.0812 Length: 2053
Score: 52.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-774-176-17 (1-9) x HS5T40A (1-2053)
Qy 1 HisMetAlaAspMetValThrTrpLeu 9
|||||
Db 985 CACATGGCAGACATGGTGACCTGGCTC 1011
|||||

RESULT 11
BD127282

LOCUS BD127282 2359 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127282
VERSION BD127282.1 GI:232222227
KEYWORDS JP 2002017375-A/2713.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2359)
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002017375-A/2713
PD 22-JAN-2002
PF 07-JUL-2000 JP 200253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUUI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
C12P21/02,C12O1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FT CDS Location/Qualifiers
1..2359
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0.096 Length: 2359
Score: 52.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-17 (1-9) x BD127282 (1-2359)
Qy 1 HisMetAlaAspMetValThrTrpLeu 9
|||||
Db 1324 CACATGGCAGACATGGTGACCTGGCTC 1350
|||||

RESULT 12
CQ782724

LOCUS CQ782724 2359 bp DNA linear PAT 17-MAR-2004
DEFINITION Sequence 2864 from Patent EP1396543.
ACCESSION CQ782724
VERSION CQ782724.1 GI:45502667
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primers for synthesizing full length cDNA clones and their use
JOURNAL Patent: EP 1396543-A 2864 10-MAR-2004;
Research Association for Biotechnology (JP)

```

FEATURES             Location/Qualifiers
     source            1..2359
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
     CDS                424..1575
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAF85958.1"
                        /db_xref="GI:45502668"
                        /translation="MPGCSRGPAAGDGLRLALVLLGWSSSSPTSSASSFSSS
                        APFLASVSAQPPPLDQPCALCESEARTVKCVNRNLTETPTDLPAYVRLFLTGNO
                        LAVLPAGAFARPPPLAELALNLSGRSLDEVRAGAFELHPSLRQLDLSHNPLADLSFF
                        AFSGSNASVSAPSLVELILNHIVPDERQNRSEFGVMVAALLAGRALQGLRLLELA
                        SNHFLYLRDVLIAQLPSLRHLDLSNNLSVLTYSFNRNLTHLESLELDNALKVLHNG
                        TLAEQLGLPHIRVFLDNNPWCDCHMADMTWLKETEYVQGKDRLTCAYPEKMRNRVL
                        LELNSADLDCDPLPSPSLQTSYVFLGIVLALIGALFLLVILNRKGIKK"

ORIGIN
Alignment Scores:
Pred. No.:          0.096      Length:          2359
Score:              52.00      Matches:          9
Percent Similarity: 100.0%      Conservative:    0
Best Local Similarity: 100.0%      Mismatches:     0
Query Match:        100.0%      Indels:         0
DB:                  2          Gaps:         0

US-10-774-176-17 (1-9) x CQ782724 (1-2359)

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      1324 CACATGGCAGACATGGTGACCTGGCTC 1350

RESULT 13
AX961916
LOCUS      Sequence 2359 bp mRNA linear PRI 20-JAN-2006
DEFINITION Homo sapiens CDNA FLJ90305 fis, clone NT2RP2000694, highly similar
            to Homo sapiens 5T4 oncofetal trophoblast glycoprotein gene.
ACCESSION  AK074786
VERSION     AK074786.1 GI:22760460
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE     Homo sapiens
            Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.

REFERENCE   1
AUTHORS    Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
            Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
            Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
            Kojima,S., Nagahara,K., Maguho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
            Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
            Ninomiya,K.
TITLE      NEDO human cDNA sequencing project
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 2359)
AUTHORS    Isogai,T. and Otsuki,T.
TITLE      Direct Submission
JOURNAL    Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:flj-cdn@nifty.com, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology; cDNA library construction:
            Institute of Medical Science, University of Tokyo, Laboratory of
            Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
            sequencing and clone selection: Helix Research Institute (supported
            by Japan Key Technology Center etc.).

FEATURES             Location/Qualifiers
     source            1..2359
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"

```

```

/clone="NT2RP2000694"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3
mRNA from NT2 neuronal precursor cells after 2-weeks
retinoic acid (RA) induction"

ORIGIN
Alignment Scores:
Pred. No.:          0.096      Length:          2359
Score:              52.00      Matches:          9
Percent Similarity: 100.0%      Conservative:    0
Best Local Similarity: 100.0%      Mismatches:     0
Query Match:        100.0%      Indels:         0
DB:                  5          Gaps:         0

US-10-774-176-17 (1-9) x AK074786 (1-2359)

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      1324 CACATGGCAGACATGGTGACCTGGCTC 1350

RESULT 14
AX961916
LOCUS      Sequence 2361 bp DNA linear PAT 14-JAN-2004
DEFINITION Stat6 activation gene
ACCESSION  AX961916
VERSION     AX961916.1 GI:40881326
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.

REFERENCE   1
AUTHORS    Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
TITLE      Stat6 activation gene
JOURNAL    Patent: WO 03104277-A 127 18-DEC-2003;
            Asahi Kasei Kabushiki Kaisha (JP)
FEATURES   source
            1..2361
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            426..1688
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAF06467.1"
            /db_xref="GI:40881327"
            /translation="MPGCSRGPAAGDGLRLALVLLGWSSSSPTSSASSFSSS
            APFLASVSAQPPPLDQPCALCESEARTVKCVNRNLTETPTDLPAYVRLFLTGNO
            LAVLPAGAFARPPPLAELALNLSGRSLDEVRAGAFELHPSLRQLDLSHNPLADLSFF
            AFSGSNASVSAPSLVELILNHIVPDERQNRSEFGVMVAALLAGRALQGLRLLELA
            SNHFLYLRDVLIAQLPSLRHLDLSNNLSVLTYSFNRNLTHLESLELDNALKVLHNG
            TLAEQLGLPHIRVFLDNNPWCDCHMADMTWLKETEYVQGKDRLTCAYPEKMRNRVL
            LELNSADLDCDPLPSPSLQTSYVFLGIVLALIGALFLLVILNRKGIKKWMHNRDACC
            RDHMEGYHYRYEINADPLRLTNSSDV"

ORIGIN
Alignment Scores:
Pred. No.:          0.0961     Length:          2361
Score:              52.00      Matches:          9
Percent Similarity: 100.0%      Conservative:    0
Best Local Similarity: 100.0%      Mismatches:     0
Query Match:        100.0%      Indels:         0
DB:                  2          Gaps:         0

US-10-774-176-17 (1-9) x AX961916 (1-2361)

Qy      1 HisMetAlaAspMetValThrTrpLeu 9
Db      1326 CACATGGCAGACATGGTGACCTGGCTC 1352

```

```

RESULT 15
BD127283          2361 bp      DNA      linear      PAT 18-SEP-2002
LOCUS
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD127283
VERSION
BD127283.1 GI:23222228
KEYWORDS
JP 2002017375-A/2714.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
1 (bases 1 to 2361)
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002017375-A 2714 22-JAN-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2714
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
Location/Qualifiers
FT CDS Location/Qualifiers
1..2361
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0.0961 Length: 2361
Score: 52.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-17 (1-9) x BD127283 (1-2361)
Qy 1 HisMetalAspMetValThrTrpLeu 9
Db 1326 CACATGGCAGACATGGTGACCTGGCTC 1352

```

Search completed: May 27, 2006, 19:34:57
 Job time : 3362.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-16

Perfect score: 59

Sequence: 1 FLDNNPWVC 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB pool/US10774176/runat_26052006_091441_24976/app_query.fasta_1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02h
-USER=US10774176 @CGN 1 1 2389 @runat_26052006_091441_24976 -NCPU=6 -ICPU=3
-NO MMAP -NRG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*
15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	59	100.0	130	10	ADK11793	Adk11793 Breast ca
2	59	100.0	299	10	ACD93536	Ac93536 Human col
3	59	100.0	453	5	AAS87174	Aas87174 DNA encod

4	59	100.0	475	13	ADU11677	Adu11677 Solid tum
5	59	100.0	901	3	AAA27060	Aaa27060 Canine 5T
6	59	100.0	927	6	ABT07721	Abt07721 Breast ca
7	59	100.0	927	8	ABX76333	Abx76333 Lung canc
8	59	100.0	927	10	ADB80503	Adb80503 Ovarian c
9	59	100.0	927	11	ADN38723	Adn38723 Cancer/an
10	59	100.0	973	8	AAD56198	Aad56198 Human LRR
11	59	100.0	1156	6	ABV99349	Abv99349 Human NOV
12	59	100.0	1260	6	ABK87175	Abk87175 cDNA enco
13	59	100.0	1260	10	ADB97513	Adb97513 Feline 5T
14	59	100.0	1260	10	ADB97452	Adb97452 DNA encod
15	59	100.0	1263	3	AAA27058	Aaa27058 Human 5T4
16	59	100.0	1263	4	AAF99736	Aaf99736 Nucleotid
17	59	100.0	1263	6	ABK87174	Abk87174 cDNA enco
18	59	100.0	1281	3	AAA27059	Aaa27059 Mouse 5T4
19	59	100.0	1331	8	AAD56199	Aad56199 Human LRR
20	59	100.0	2020	10	ADJ56299	Adj56299 Human CDN
21	59	100.0	2053	8	ACC51052	Acc51052 Human bla
22	59	100.0	2053	8	ABX76332	Abx76332 Lung canc
23	59	100.0	2053	8	AAD56197	Aad56197 Human LRR
24	59	100.0	2053	8	AAD56200	Aad56200 Human LRR
25	59	100.0	2053	11	ADN38721	Adn38721 Cancer/an
26	59	100.0	2053	12	ADL06473	Adl06473 Human tum
27	59	100.0	2053	12	ADN03961	Adn03961 Antipsori
28	59	100.0	2053	13	ADR25444	Adr25444 Breast ca
29	59	100.0	2053	13	ACN38510	Acn38510 Tumour-as
30	59	100.0	2053	13	ADV35098	Adv35098 Human CDN
31	59	100.0	2053	14	AED17761	Aed17761 Fibrotic
32	59	100.0	2338	5	AAS87175	Aas87175 DNA encod
33	59	100.0	2359	4	AAK94253	Aak94253 Human ful
34	59	100.0	2359	12	ADL30831	Adl30831 Full leng
35	59	100.0	2361	4	AAK94254	Aak94254 Human ful
36	59	100.0	2361	12	ADI26162	Adi26162 Human CDN
37	59	100.0	2361	12	ADL30833	Adl30833 Full leng
38	59	100.0	2557	12	ADI26160	Adi26160 Human CDN
39	59	100.0	2557	12	ADI26158	Adi26158 Human CDN
40	47	79.7	2710	11	ADM02391	Adm02391 Human CDN
41	47	79.7	2710	14	AEC85321	Aec85321 Human CDN
42	46	78.0	286	6	ABK78385	Abk78385 Bacillus
43	46	78.0	286	6	ABK78397	Abk78397 Bacillus
44	46	78.0	28564	10	ADD46508	Add46508 Human gen
45	45	76.3	561	13	ADX27350	Adx27350 Plant ful

ALIGNMENTS

RESULT 1
ADK11793
ID ADK11793 standard; DNA; 130 BP.
AC ADK11793;
XX
XX
DT 06-MAY-2004 (first entry)
XX
DE Breast cancer differentially expressed gene product #199.
XX
KW ds; cytostatic; gene therapy; DKFZp5661133 activity inhibitor;
XX
KW breast cancer; differential expression.
XX
OS Homo sapiens.
XX
PN WO2003057925-A1.
XX
PD 17-JUL-2003.
XX
PF 08-JAN-2003; 2003WO-US0000657.
XX
PR 08-JAN-2002; 2002US-0345637P.
XX
XX (CHIR) CHIRON CORP.
XX
PI Hansen R;
XX

DR WPI; 2003-577534/54.
 XX Inhibiting a cancerous phenotype of a cell, useful for treating breast
 PT cancer comprises contacting a cancerous mammalian cell with an agent for
 PT inhibition of DKFZp5661133 activity.
 XX
 XX Claim 30; SEQ ID NO 199; 257pp; English.
 XX
 CC The invention relates to a method of inhibiting a cancerous phenotype of
 CC a cell comprises contacting a cancerous mammalian cell with an agent for
 CC inhibition of DKFZp5661133 activity. The methods are useful for treating
 CC cancer, e.g. breast cancer. This sequence represents a gene product which
 CC is differentially expressed in breast cancer cells. The sequence can be
 CC used in the method of the invention.
 XX
 SQ Sequence 130 BP; 30 A; 41 C; 30 G; 29 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.0973 Length: 130
 Score: 59.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-16 (1-9) x ADK11793 (1-130)
 Qy 1 PheLeuAspAenAenProTrpValCys 9
 Db 90 TTCTGGACAAATCCCTGGGTCTGC 116
 RESULT 2
 ACD93536
 ID ACD93536 standard; cDNA; 299 BP.
 XX
 AC ACD93536;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human colon cancer cell expressed cDNA #1948.
 XX
 KW Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KW agriculture; food crop genome; resistance gene; retrovirus;
 KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002155438-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 27-SEP-1999; 99US-00406117.
 XX
 XX 20-NOV-1998; 98US-00196716.
 PR
 XX (SIMP/) SIMPSON A J G.
 PA (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 XX
 XX Simpson AJG, Neto ED, Brentani RR;
 XX
 XX WPI; 2003-182626/18.
 DR
 XX Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 XX
 XX Example 9; Page 302; 959pp; English.
 PS
 XX

CC The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (I) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a foetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 XX
 SQ Sequence 299 BP; 75 A; 84 C; 78 G; 62 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.248 Length: 299
 Score: 59.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-16 (1-9) x ACD93536 (1-299)
 Qy 1 PheLeuAspAenAenProTrpValCys 9
 Db 162 TTCTGGACAAATCCCTGGGTCTGC 188
 RESULT 3
 AAS87174
 ID AAS87174 standard; cDNA; 453 BP.
 XX
 AC AAS87174;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #22978.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2001/5067-A2.
 XX
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 PR

```
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG22987.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 22978; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 453 BP; 108 A; 111 C; 113 G; 121 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 0.396 Length: 453
XX Score: 59.00 Matches: 9
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 100.0% Indels: 0
XX DB: 5 Gaps: 0
XX
XX US-10-774-176-16 (1-9) x AAS87174 (1-453)
XX
XX Qy 1 PheLeuAspAsnAsnProTrpValCys 9
XX Db 55 TTCCTGGACAAACATCCCTGGGTCTGC 81
XX
XX RESULT 4
XX ADU11677
XX ID ADU11677 standard; DNA; 475 BP.
XX AC ADU11677;
XX XX ADU11677;
XX DT 27-JAN-2005 (first entry)
XX XX
XX DE Solid tumour prognosis gene seqid 2116.
XX KW cytostatic; gene therapy; expression profile; solid tumour;
XX KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX XX Unidentified.
XX OS
XX XX WO2004097052-A2.
XX FN
XX XX 11-NOV-2004.
XX PD
XX XX 29-APR-2004; 2004WO-US013587.
XX PF
XX
XX 29-APR-2003; 2003US-0466067P.
XX 23-JAN-2004; 2004US-0538246P.
XX (AMHP ) WYETH.
XX (STRA/) STRAHS A.
XX
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
XX Immerman F, Dorner AJ;
XX WPI; 2004-804779/79.
XX
XX A method, useful for prognosing and treating solid tumor, comprises
XX comparing an expression profile of a gene expressed in peripheral blood
XX mononuclear cells to a reference expression profile of a gene.
XX
XX Disclosure; Page; 111pp; English.
XX
XX The invention describes a method comprising comparing an expression
XX profile of at least one gene in a peripheral blood sample of a patient to
XX at least one reference expression profile of the at least one gene, where
XX the patient has a solid tumour, and each of the gene is differentially
XX expressed in peripheral blood mononuclear cells (PBMCs) of a first class
XX of patients as compared to PBMCs of a second class of patients, where
XX both the first and second classes of patients have the solid tumour, and
XX each of the first and second classes is a subcluster formed by an
XX unsupervised clustering analysis of gene expression profiles in PBMCs of
XX a population of patients who have the solid tumour, and where the
XX majority of the first class of patients has a first clinical outcome, and
XX the majority of the second class of patients has a second clinical
XX outcome. Also described are: a system comprising (i) a memory or a
XX storage medium including data that represent an expression profile of at
XX least one gene in a peripheral blood sample of a patient who has a solid
XX tumour, (ii) at least another storage medium including data that
XX represent at least one reference expression profile of the gene, (iii) a
XX program capable of comparing the expression profile to the reference
XX expression profile, and (iv) a processor capable of executing the
XX program, where expression levels of the gene in peripheral blood
XX mononuclear cells of patients who have the solid tumour correlate with
XX clinical outcomes of the patients; and a nucleic acid or protein array
XX comprising concentrated probes for solid tumour prognosis genes, where
XX each of the solid tumour prognosis genes is differentially expressed in
XX PBMCs of a first class of patients as compared to PBMCs of a second class
XX of patients, where both the first and second classes of patients have a
XX solid tumour, and where the first class of patients has a first clinical
XX outcome, and the second class of patients has a second clinical outcome.
XX The method, system, and array are useful for prognosing and treating
XX solid tumours. This sequence represents a solid tumour prognosis gene of
XX the invention. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 0.418 Length: 475
XX Score: 59.00 Matches: 9
XX Percent Similarity: 100.0% Conservative: 0
XX Best Local Similarity: 100.0% Mismatches: 0
XX Query Match: 100.0% Indels: 0
XX DB: 13 Gaps: 0
XX
XX US-10-774-176-16 (1-9) x ADU11677 (1-475)
XX
XX Qy 1 PheLeuAspAsnAsnProTrpValCys 9
XX Db 147 TTCCTGGACAAACATCCCTGGGTCTGC 173
XX
XX RESULT 5
XX AAA27060
XX ID AAA27060 standard; DNA; 901 BP.
XX XX
XX AC AAA27060;
```



```

PR 09-APR-2001; 2001US-0282698P.
PR 04-MAY-2001; 2001US-0288590P.
PR 29-MAY-2001; 2001US-029443P.
XX
PR (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PR Mack DH, Gish KC, Afar D;
XX
DR WPI; 2002-583738/62.
DR N-PSDB; ABU05564.
XX
PT Detecting a breast cancer-associated transcript in a patient's cell,
PT useful for diagnosing breast cancer, comprises contacting a biological
PT sample with a polynucleotide that selectively hybridizes with breast
PT cancer nucleic acids.
XX
PS Claim 9; Page 372; 414pp; English.
XX
CC The invention comprises a method of detecting a breast cancer-associated
CC transcript in a cell from a patient. The method of the invention involves
CC contacting a biological sample from the patient with a nucleotide that
CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
CC in the specification. The method of the invention is useful in the
CC diagnosis or prognosis of breast cancer, and for detecting genes that are
CC up or down-regulated in breast cancer cells. Genes identified by the
CC method of the invention can be used in diagnostic purposes and also as
CC targets for screening for therapeutic compounds that modulate breast
CC cancer (e.g. hormones or antibodies). Identification of genes that are
CC over or under expressed in breast cancer can additionally provide high-
CC resolution, high-sensitivity datasets which can be used in the areas of
CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
CC represent the 69 breast cancer-associated gene sequences of the invention
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.886 Length: 927
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-16 (1-9) x ABT07721 (1-927)
QY 1 PheLeuAspAsnProTrpValCys 9
Db 526 TTCCTGGACAAACAATCCCTGGGTCTGC 552
RESULT 7
ABX76333
ID ABX76333 standard; DNA; 927 BP.
XX
AC ABX76333;
XX
XX
DT 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polynucleotide #197.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
XX WO200286443-A2.
PN
XX 31-OCT-2002.
PD
XX 18-APR-2002; 2002WO-US012476.
PF

18-APR-2001; 2001US-0284770P.
10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.
PR 13-NOV-2001; 2001US-0350666P.
PR 29-NOV-2001; 2001US-0334370P.
PR 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Aziz N, Murray R;
XX
XX WPI; 2003-093161/08.
XX
XX P-PSDB; ABU56604.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 22; Page 336; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
CC invention
XX
SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.886 Length: 927
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-16 (1-9) x ABX76333 (1-927)
QY 1 PheLeuAspAsnProTrpValCys 9
Db 526 TTCCTGGACAAACAATCCCTGGGTCTGC 552
RESULT 8
ADB80503
ID ADB80503 standard; DNA; 927 BP.
XX
XX ADB80503;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Ovarian cancer-associated transcript #34.
DE
XX
XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection; ds; gene.
XX
XX Homo sapiens.
OS

```

XX Key Location/Qualifiers
 FH 1. 927
 FT /*tag= a
 FT
 XX WO2002102235-A2.
 PN
 XX 27-DEC-2002.
 PD
 XX 18-JUN-2002; 2002WO-US019297.
 XX
 PF 18-JUN-2001; 2001US-02999234P.
 PR 27-AUG-2001; 2001US-0315287P.
 PR 05-SEP-2001; 2001US-0317544P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Mack DH, Gish KC;
 XX
 XX WPI; 2003-167431/16.
 DR P-PSDB; ADB80504.
 XX
 XX Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.
 XX
 PS Claim 10; Page 297; 332pp; English.
 XX
 CC The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the nucleic acids
 CC used for the detection method of the invention.
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.886 Length: 927
 Score: 59.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-16 (1-9) x ADB80503 (1-927)
 QY 1 PheLeuAspAenAanProTrpValCys 9
 Db 526 TTCCTGGACACAAATCCCTGGTCTGC 552
 RESULT 9
 ADN38723
 ID ADN38723 standard; cDNA; 927 BP.
 XX
 AC ADN38723;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.
 XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine; gene; ss.
 XX Homo sapiens.
 OS
 XX WO2003042661-A2.
 PN
 XX 22-MAY-2003.
 PD
 XX 13-NOV-2002; 2002WO-US036810.
 XX
 PF 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 PI
 XX WPI; 2003-468649/44.
 DR P-PSDB; ADN38724.
 XX
 DR Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO 41; 1385pp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0.886 Length: 927
 Score: 59.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0

```
DB: 11 Gaps: 0
US-10-774-176-16 (1-9) x ADN38723 (1-927)
Qy 1 PheLeuAspAsnProTrpValCys 9
Db 536 TTCCTGGACAACAATCCCTGGGTCTGC 552

RESULT 10
AAD56198
ID AAD56198 standard; DNA; 973 BP.
AC AAD56198;
XX
DT 07-AUG-2003 (first entry)
XX
DE Human LRRCAPS related DNA #5.
XX
KW Human; p53 pathway; Leucine rich repeat capricious related protein;
KW LRRCAPS; cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO2003035831-A2.
XX
PD 01-MAY-2003.
XX
PF 21-OCT-2002; 2002WO-US033540.
XX
PR 22-OCT-2001; 2001US-0338733P.
PR 15-FEB-2002; 2002US-0357600P.
PR 01-MAR-2002; 2002US-0361196P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
PI Francis-Lang H, Friedman L;
XX
DR WPI; 2003-421410/39.
XX
PT Identifying a candidate p53 pathway-modulating agent for treating cancer
PT comprises contacting an assay system comprising a purified leucine rich
PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX
PS Example 5; Page 74-75; 99pp; English.
XX
CC The invention relates to a method of identifying a candidate p53 pathway
CC modulating agent. The method involves contacting an assay system
CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
CC polypeptide or nucleic acid or its fragment with a test agent and
CC detecting a test agent-biased activity, where a difference between the
CC test agent-biased activity and the reference activity identifies the test
CC agent as a candidate p53 pathway modulating agent. The method is useful
CC for identifying a candidate p53 pathway-modulating agent for preparing a
CC composition for diagnosing or treating cancer. The invention is useful in
CC gene therapy. The present sequence is human LRRCAPS related DNA
XX
SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.935 Length: 973
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-16 (1-9) x AAD56198 (1-973)
Qy 1 PheLeuAspAsnProTrpValCys 9
Db 541 TTCCTGGACAACAATCCCTGGGTCTGC 567
```

```
RESULT 11
ABV99349
ID ABV99349 standard; DNA; 1156 BP.
XX
AC ABV99349;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human NOV8a coding sequence.
XX
KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antifertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX
OS Homo sapiens.
XX
PN WO200272771-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002WO-US007288.
XX
PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-027676P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
```

PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 04-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX (CURA-) CURAGEN CORP.
XX
XX
PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
PI Boldog FL, Li L, Zerhusen BD, Tchervnev VT, Gangolli EA, Vernet CAM;
PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
PI Voss EZ, Malyankar UM, Anderson DM, Patturajan M, Miller CE;
PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VV, Pochart PF;
PI Zhong M;
XX
XX WPI; 2002-732824/79.
DR P-PSDB; ABP70071.
XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
PT disorders, and asthma.
XX
XX Claim 16; Page 114-115; 619pp; English.
XX
XX The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, haematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX
SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.14 Length: 1156
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-774-176-16 (1-9) x ABV99349 (1-1156)
QY 1 PheLeuAspAenAenProTirpValCys 9
|||||

Db 757 TTCCTGGACACAAATCCCTGGGTCTGC 783
RESULT 12
ABK87175
ID ABK87175 standard; cDNA; 1260 BP.
XX
AC ABK87175;
XX
DT 07-OCT-2002 (first entry)
XX
DE cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
XX
KW Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
KW cell proliferative disorder; infection; inflammatory condition;
KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
KW foetal abnormality; foetal sex determination; gene; ss.
XX
OS Fells sp.
XX
FH Key Location/Qualifiers
CDS 1..1260
FT /*tag= a
FT /product= "5T4 protein"
XX
XX WO200238612-A2.
XX
XX 16-MAY-2002.
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Myers K, Drury N, Carroll M;
PI
XX WPI; 2002-557449/59.
DR P-PSDB; AAU98694.
XX
XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
PT polypeptide, useful in preparation of vaccine for treating and/or
PT preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 4; Page 68; 68pp; English.
XX
XX The present invention relates to the isolation of canine and feline
CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
CC a significant proportion of tumours. The sequences of the invention are
CC useful in a pharmaceutical composition for the prevention and/or
CC treatment of tumours or other diseases associated with cell
CC proliferation, infections, and inflammatory conditions in animals,
CC preferably dogs or cats. The compositions may also be used for cancer
CC immunotherapy in these animals. The sequences of the invention may also
CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
CC measurement and localisation of 5T4 in extracts of plasma, urine,
CC tissues, and in cell culture media. Antibodies specific for the 5T4
CC protein are useful for isolating foetal cells from maternal blood. The
CC isolation process may form part of a diagnostic method e.g. the foetal
CC cells may then be subject to biochemical or genetic sampling used for
CC testing foetal abnormalities, or to determine the sex of the foetus(es).
CC The present sequence encodes feline 5T4 protein
XX
SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 1.25 Length: 1260
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-16 (1-9) x ABK97175 (1-1260)

Qy 1 PheLeuAspAsnAsnProTrpValCys 9
 Db 865 TTCTCGACAAACATCCCTGGGTCTGC 891

RESULT 13
 ADB97513
 ID ADB97513 standard; DNA; 1260 BP.
 XX
 XX ADB97513;
 XX AC
 XX DT 04-DEC-2003 (first entry)
 XX
 XX DE Feline 574 antigen DNA.
 XX
 XX KW Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 574 antigen; 574 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.
 XX
 XX OS Unidentified.
 XX
 XX FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 574 antigen protein"
 XX
 XX PN WO2003068816-A1.
 XX
 XX PD 21-AUG-2003.
 XX
 XX PF 13-FEB-2003; 2003WO-GB000670.
 XX
 XX PR 13-FEB-2002; 2002GB-00003419.
 XX
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 XX PI Carroll M, Kingman S, Redchenko I;
 XX WPI; 2003-637141/60.
 XX P-PSDB; ADB97520.
 XX
 XX PT New major histocompatibility complex class I peptide epitopes from human
 PT 574 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.
 XX
 XX PS Disclosure; Page 67; 73pp; English.
 XX
 XX CC The invention relates to a novel Major Histocompatibility Complex (MHC)
 CC class I peptide epitope of the 574 antigen. The invention further
 CC provides a polypeptide string comprising the 574 epitope; a nucleic acid
 CC sequence encoding the 574 epitope or a polypeptide string of the 574
 CC epitope; a vector system capable of delivering the 574 epitope nucleic
 CC acid to a cell; a cell pulsed with the 574 epitope, a polypeptide of the
 CC 574 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 574 epitope and/its encoding nucleic acid; a method
 CC comprising detecting the presence of the 574 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 574 epitope in conjunction with an MHC class
 CC I molecule. The 574 epitope has cytostatic activity. The vaccine
 CC comprising the 574 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 574 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline
 CC 574 antigen coding DNA of the invention.

Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1.25 Length: 1260
 Score: 59.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0

US-10-774-176-16 (1-9) x ADB97513 (1-1260)

Qy 1 PheLeuAspAsnAsnProTrpValCys 9
 Db 865 TTCTCGACAAACATCCCTGGGTCTGC 891

RESULT 14
 ADB97452
 ID ADB97452 standard; DNA; 1260 BP.
 XX
 XX AC ADB97452;
 XX
 XX DT 04-DEC-2003 (first entry)
 XX
 XX DE DNA encoding feline 574 protein.
 XX
 XX KW gene; ds; feline; Major Histocompatibility Complex class II; MHC;
 KW epitope; 574 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
 XX
 XX OS Unidentified.
 XX
 XX FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 574 antigen protein"
 XX
 XX PN WO2003068815-A2.
 XX
 XX PD 21-AUG-2003.
 XX
 XX PF 13-FEB-2003; 2003WO-GB000618.
 XX
 XX PR 13-FEB-2002; 2002GB-00003420.
 XX
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 XX PI Carroll M, Harrop R, Kingsman S;
 XX WPI; 2003-663795/62.
 XX P-PSDB; ADB97455.
 XX
 XX PT New Major Histocompatibility Complex class II peptide epitope of 574,
 PT useful for manufacturing a medicament for diagnosing, preventing and/or
 PT treating a disease, e.g. cancer.
 XX
 XX PS Disclosure; Page 49; 63pp; English.
 XX
 XX CC The invention relates to a Major Histocompatibility Complex (MHC) class
 CC II peptide epitope of the 574 antigen. The vaccine or T-cell line or
 CC clone has a cytostatic activity, as it is useful in manufacturing a
 CC medicament for preventing and/or treating a disease, particularly cancer.
 CC The methods are useful for detecting T-cells capable of specifically
 CC recognising a peptide epitope in conjunction with an MHC molecule, for
 CC diagnosing or monitoring the progression of a cancerous disease, or for
 CC detecting the presence of a peptide or nucleic acid using an agent. The
 CC MHC class II peptide epitope of the invention can be used in gene therapy
 CC or as part of a vaccine. This polynucleotide sequence represents the DNA
 CC coding for the feline 574 protein.

Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1.25 Length: 1260
 Score: 59.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0

```

DB:
2 Gaps: 0
US-10-774-176-16 (1-9) x AX316088 (1-901)

QY 1 PheLeuAspAsnAsnProTrpValCys 9
|||||
Db 383 TTCTGGACAAACACCCCTGGGTCTGC 409

RESULT 5
AX829164
LOCUS AX829164
DEFINITION Sequence 57 from Patent WO02059377.
ACCESSION AX829164
VERSION AX829164.1 GI:39838931
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 Mack,D.H., Gish,K.C. and Afar,D.
Methods of diagnosis of breast cancer, compositions and methods of
screening for modulators of breast cancer
JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
EOS Biotechnology, Inc. (US)
FEATURES
source
1..927
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0.0295 Length: 927
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-16 (1-9) x AX829164 (1-927)

QY 1 PheLeuAspAsnAsnProTrpValCys 9
|||||
Db 526 TTCTGGACAAACATCCCTGGGTCTGC 552

RESULT 6
DD161112
LOCUS DD161112
DEFINITION Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
Encoding The Antigens, and Methods of Use.
ACCESSION DD161112
VERSION DD161112.1 GI:83967439
KEYWORDS JP 2005508604-A/23.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1156)
Padigaru,M., Shenoy,S.G., Pochart,P.F., Kekuda,R., Gusev,V.Y.,
Zhong,M., Jr,R.J.T., Casman,S.J., Li,L., Miller,C.E.,
Patturajan,M., Anderson,D.W., Malyankar,U.M., Voss,E.Z.,
Spaderna,S.K., Gorman,L., Spytek,K.A., Liu,X., Burgess,C.B.,
Pena,C.E.A., Gerlach,V., Smithson,G., Mezes,P.D., Rastelli,L.,
Boldog,F.L., Guo,X., Vernet,C.A.M., Gangolli,E.A., Tchernev,V.T.
and Zerkhusen,B.D.
Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
Encoding The Antigens, and Methods of Use
JOURNAL Patent: JP 2005508604-A 23 07-APR-2005;
Muralidhara Padigaru,Suresh Shenoy,Remesh Kekuda,Vladimir Gusev,
Pascale Pochart,Mei Zhong,Luca Rastelli,Peter Mezes, Glennada

```

```

Smithson, Xiaojia Guo, Valerie Gerlach, Stacie Casman, Ferenc
Boldog, Li Li, Bryan Zerkhusen, Velizar Tchernev, Esha Gangolli, Corine
Vernet, Carol Pena, Catherine Burgess, Xiaohong Liu, Kimberly
Spytek, Linda Gorman, Steven Spaderna, Edward Voss, Uriel
Malyankar, David Anderson, Meera Patturajan, Charles Miller, Raymond J
Taupier Jr
OS Homo sapiens
PN JP 2005508604-A/23
PD 07-APR-2005
PF 08-MAR-2002 JP 2002571827
PR 19-JUN-2001 US 60/299310, 18-JUN-2001 US 60/299027, PR
31-MAY-2001 US 60/294889, 31-MAY-2001 US 60/294899, PR
30-MAY-2001 US 60/294485, 09-MAR-2001 US 60/274849, PR
13-MAR-2001 US 60/275579, 03-MAY-2001 US 60/288342, PR
02-MAY-2001 US 60/288066, 30-APR-2001 US 60/287424, PR
13-APR-2001 US 60/283675, 04-APR-2001 US 60/281194, PR
02-APR-2001 US 60/280802, 02-APR-2001 US 60/280822, PR
03-MAY-2001 US 60/288528, 15-MAY-2001 US 60/291190, PR
16-MAY-2001 US 60/291240, 16-MAY-2001 US 60/291099, PR
12-MAR-2001 US 60/275235, 08-MAR-2001 US 60/274101, PR
08-MAR-2001 US 60/274281, 08-MAR-2001 US 60/274322, PR
08-MAR-2001 US 60/274194, 02-APR-2001 US 60/280300, PR
30-MAR-2001 US 60/280233, 30-MAR-2001 US 60/277995, PR
20-MAR-2001 US 60/277327, 20-MAR-2001 US 60/277338, PR
19-MAR-2001 US 60/276994, 16-MAR-2001 US 60/276776, PR
14-MAR-2001 US 60/276000, 13-MAR-2001 US 60/275601, PR
13-MAR-2001 US 60/275578, 20-MAR-2001 US 60/277239, PR
20-MAR-2001 US 60/277321, 21-MAR-2001 US 60/277791, PR
22-MAR-2001 US 60/277833, 23-MAR-2001 US 60/278152, PR
26-MAR-2001 US 60/278894, 27-MAR-2001 US 60/278999, PR
27-MAR-2001 US 60/279036, 28-MAR-2001 US 60/279344, PR
19-JUN-2001 US 60/299303, 10-JUL-2001 US 60/304354, PR
31-JUL-2001 US 60/309198, 03-DEC-2001 US 60/337426, PR
03-DEC-2001 US 60/338092, 21-NOV-2001 US 60/332094, PR
14-NOV-2001 US 60/33272, 14-NOV-2001 US 60/332271, PR
14-NOV-2001 US 60/332272, 14-NOV-2001 US 60/332172, PR
18-OCT-2001 US 60/330380, 27-SEP-2001 US 60/335301, PR
27-SEP-2001 US 60/325681, 12-SEP-2001 US 60/325430, PR
10-SEP-2001 US 60/318462, 03-JAN-2002 US 60/318770, PR
04-DEC-2001 US 60/337185, 08-MAR-2002 US 60/345705, PR
16-AUG-2001 US 60/312903
PI muralidhara padigaru, suresh g shenoy, pascale f-g pochart, PI
remesh kekuda,
PI vladimir y gusev, mei zhong, raymond j taupier jr, stacie j PI
casman, li li,
PI charles e miller, meera patturajan, david w anderson, uriel m PI
malyankar,
PI edward z voss, steven k spaderna, linda gorman, kimberly PI a
spytek,
PI xiaohong liu, catherine e burgess, carol e a pena, valerie PI
gerlach,
PI glennada smithson, peter d mezes, luca rastelli, ferenc l boldog,
PI xiaoja guo,
PI corine a m vernet, esha a gangolli, velizar t tchernev, bryan d
PI zerkhusen
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1156
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0.0384 Length: 1156
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

```

```

US-10-774-176-16 (1-9) x DD161112 (1-1156)
QY 1 PheLeuAspAsnAsnProTirpValCys 9
Db 757 TTCTGGACAAACATCCCTGGGTCTGC 783

RESULT 7
AX467373 1260 bp DNA linear PAT 16-JUL-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Myers,K., Drury,N. and Carroll,M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 3 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
Location/Qualifiers
1..1260
/organism="Felis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9687"
ORIGIN
Alignment Scores:
Pred. No.: 0.0425 Length: 1260
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-16 (1-9) x AX467373 (1-1260)
QY 1 PheLeuAspAsnAsnProTirpValCys 9
Db 865 TTCTGGACAAACATCCCTGGGTCTGC 891

RESULT 8
AX821533 1260 bp DNA linear PAT 10-DEC-2003
LOCUS
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
TITLE MHC class I peptide epitopes from the human 5t4 tumor-associated antigen
JOURNAL Patent: WO 03068816-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
Location/Qualifiers
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 0.0425 Length: 1260
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-16 (1-9) x AX821533 (1-1260)
QY 1 PheLeuAspAsnAsnProTirpValCys 9
Db 865 TTCTGGACAAACATCCCTGGGTCTGC 891

RESULT 9
AX821548 1260 bp DNA linear PAT 10-DEC-2003
LOCUS
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
TITLE MHC class II peptide epitope of 5t4 antigen
JOURNAL Patent: WO 03068815-A 1 21-AUG-2003;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
Location/Qualifiers
1..1260
/organism="Felis catus"
/mol_type="unassigned DNA"
/db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 0.0425 Length: 1260
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-16 (1-9) x AX821548 (1-1260)
QY 1 PheLeuAspAsnAsnProTirpValCys 9
Db 865 TTCTGGACAAACATCCCTGGGTCTGC 891

RESULT 10
BD249731 1263 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 1 17-SEP-2002;
OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS

```

PC	C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
PC	C07K19/00,
PC	C12N15/00
CC	Polypeptide
FH	Key
FT	source
FT	1. 1263
FT	Location/Qualifiers
FT	1. 1263
FT	/organism='Homo sapiens (human)'
FEATURES	
source	Location/Qualifiers
	1. .1263
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
Alignment Scores:	
Pred. No.:	0.0426
Score:	59.00
Percent Similarity:	100.0%
Best Local Similarity:	100.0%
Query Match:	100.0%
DB:	2
US-10-774-176-16 (1-9) x BD249731 (1-1263)	
QY	1 PhleuaspAsnAnProTtpValCys 9
Db	868 TTCTGGACAACAATCCCTGGGTCTGC 894
RESULT 11	
AX025011	
LOCUS	AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION	Sequence 1 from Patent WO0029428.
ACCESSION	AX025011
VERSION	AX025011.1 GI:10184932
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.
REFERENCE	1
AUTHORS	Carroll,M.W. and Myers,K.A.
TITLE	Polypeptide
JOURNAL	Patent: WO 0029428-A 1 25-MAY-2000;
	CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
	BIOMEDICA LTD (GB)
FEATURES	
source	Location/Qualifiers
	1. .1263
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
ORIGIN	
Alignment Scores:	
Pred. No.:	0.0426
Score:	59.00
Percent Similarity:	100.0%
Best Local Similarity:	100.0%
Query Match:	100.0%
DB:	2
US-10-774-176-16 (1-9) x AX025011 (1-1263)	
QY	1 PhleuaspAsnAnProTtpValCys 9
Db	868 TTCTGGACAACAATCCCTGGGTCTGC 894
RESULT 12	
AX149553	
LOCUS	AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION	Sequence 14 from Patent WO0136486.
ACCESSION	AX149553
VERSION	AX149553.1 GI:14347991

KEYWORDS	synthetic construct
SOURCE	synthetic construct
ORGANISM	other sequences; artificial sequences.
REFERENCE	1
AUTHORS	Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W., Ellard,F.M. and Myers,K.A.
TITLE	Antibodies
JOURNAL	Patent: WO 0136486-A 14 25-MAY-2001; Oxford Biomedica (UK) Limited (GB)
FEATURES	Location/Qualifiers
source	1..1263 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="574"
ORIGIN	
Alignment Scores:	Length: 1263
Pred. No.:	Score: 59.00 Matches: 9
Score:	Percent Similarity: 100.0% Conservative: 0
Best Local Similarity:	Mismatches: 0
Query Match:	Indels: 0
DB:	Gaps: 0
US-10-774-176-16 (1-9) x AX149553 (1-1263)	
Qy	1 PheLeuAspAsnAsnProTrpValCys 9
Db	868 TTCTGGACAAACACCCCTGGGTCTGC 894
RESULT 13	
AX316086	AX316086 1263 bp DNA linear PAT 14-DEC-2001
LOCUS	Sequence 1 from Patent EP1160323.
DEFINITION	AX316086
ACCESSION	AX316086.1 GI:17899278
VERSION	
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Carroll,M.W. and Myers,K.A.
TITLE	574 tumour-associated antigen for use in tumour immunotherapy
JOURNAL	Patent: EP 1160323-A 1 05-DEC-2001; Oxford Biomedica (UK) Limited (GB)
FEATURES	Location/Qualifiers
source	1..1263 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	
Alignment Scores:	Length: 1263
Pred. No.:	Score: 59.00 Matches: 9
Score:	Percent Similarity: 100.0% Conservative: 0
Best Local Similarity:	Mismatches: 0
Query Match:	Indels: 0
DB:	Gaps: 0
US-10-774-176-16 (1-9) x AX316086 (1-1263)	
Qy	1 PheLeuAspAsnAsnProTrpValCys 9
Db	868 TTCTGGACAAACAACTCCCTGGGTCTGC 894
RESULT 14	
AX467371	AX467371 1263 bp DNA linear PAT 16-JUL-2002
LOCUS	

DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX467371 GI:21900602
VERSION AX467371.1
KEYWORDS
SOURCE
ORGANISM Canis sp.
Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE 1
AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. .1263
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN

Alignment Scores:
Pred. No.: 0.0426 Length: 1263
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-16 (1-9) x AX467371 (1-1263)

Qy 1 PheLeuAspAsnAsnProTrrpValCys 9
Db 868 TTCTGGACACACACCCCTGGTCTGC 894

RESULT 15

BD249732
LOCUS 1281 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249732
VERSION BD249732.1 GI:33059502
KEYWORDS JP 2002530060-A/2.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1281)
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE Polypeptide
JOURNAL Patent: JP 2002530060-A 2 17-SEP-2002;
OXFORD BIOMEDICA LTD

COMMENT
OS Mus musculus (mouse)
PN JP 2002530060-A/2
PD 17-SEP-2002
PF 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2, 27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL, KEVIN ALAN MYERS
PC C12N15/09, A61K39/00, A61K48/00, A61P35/00, C07K7/06, C07K14/065,
PC C07K19/00,
PC C12N15/00
CC Polypeptide
FH Key
FT source
1. .1281
Location/Qualifiers
/organism="Mus musculus (mouse)".

FEATURES
source
1. .1281
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 0.0433 Length: 1281
Score: 59.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-16 (1-9) x BD249732 (1-1281)
Qy 1 PheLeuAspAsnAsnProTrrpValCys 9
Db 886 TTCTGGACACACACCCCTGGTCTGC 912
Search completed: May 27, 2006, 19:35:22
Job time : 3359.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 27, 2006, 09:34:35 ; Search time 377.5 Seconds
(without alignments)
249.339 Million cell updates/sec

Title: US-10-774-176-15
Perfect score: 40
Sequence: 1 LIGAIFLLV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB pool/US10774176/runat 26052006.091441.24976/app query.fasta 1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0-
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02h
-USER=US10774176 @CGN 1.1.2389 @runat 26052006.091441.24976 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2002as:*
6: geneseqn2002bs:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2003ds:*
11: geneseqn2004as:*
12: geneseqn2004bs:*
13: geneseqn2005s:*
14: geneseqn2005bs:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	246	10	ADK11641 Breast ca
2	40	100.0	475	13	ADU11677 Solid tum
3	40	100.0	901	3	AAA27060 Canine 5T

4	40	100.0	927	6	ABT07721 Breast ca
5	40	100.0	927	8	ABX76333 Lung canc
6	40	100.0	927	10	ADB80503 Ovarian c
7	40	100.0	927	11	ADN38723 Cancer/an
8	40	100.0	973	8	AAD56198 Human LRR
9	40	100.0	1156	6	ABV99349 Human NOV
10	40	100.0	1260	6	ABK87175 cDNA enco
11	40	100.0	1260	10	ADB97513 Feline 5T
12	40	100.0	1260	10	ADB97452 DNA enco
13	40	100.0	1263	3	AAA27058 Human 5T4
14	40	100.0	1263	4	AAF89736 Nucleotid
15	40	100.0	1263	6	ABK87174 cDNA enco
16	40	100.0	1281	3	AAA27059 Mouse 5T4
17	40	100.0	1331	8	AAD56199 Human LRR
18	40	100.0	2020	10	ADJ56299 Human CDN
19	40	100.0	2053	8	ACC51052 Human bla
20	40	100.0	2053	8	ABX76332 Lung canc
21	40	100.0	2053	8	AAD56197 Human LRR
22	40	100.0	2053	8	AAD56200 Human LRR
23	40	100.0	2053	11	ADN38721 Cancer/an
24	40	100.0	2053	12	ADL06473 Human tum
25	40	100.0	2053	12	ADN03961 Antipori
26	40	100.0	2053	13	ADR25444 Breast ca
27	40	100.0	2053	13	ACN38510 Tumour-as
28	40	100.0	2053	13	ADV35098 Human CDN
29	40	100.0	2053	14	AED17761 Fibrotic
30	40	100.0	2338	5	AAS87175 DNA enco
31	40	100.0	2359	4	AAK94253 Human ful
32	40	100.0	2359	12	ADL30831 Full leng
33	40	100.0	2361	4	AAK94254 Human ful
34	40	100.0	2361	12	ADI26162 Human CDN
35	40	100.0	2361	12	ADL30833 Full leng
36	40	100.0	2557	12	ADI26160 Human CDN
37	40	100.0	2557	12	ADI26158 Human CDN
38	37	92.5	7574	2	AAI33089 Enterococ
39	37	92.5	7574	6	ABS98884 Enterococ
40	36	90.0	922	4	AAC92268 Green flu
41	36	90.0	922	9	AAL57769 Aequorea
42	36	90.0	1182	14	AED00392 Lactobaci
43	36	90.0	1242	13	ADU25709 cDNA enco
44	36	90.0	1242	14	AED00129 Lactobaci
45	36	90.0	1467	4	AAS53033 Enterococ

ALIGNMENTS

RESULT 1
ADK11641
ID ADK11641 standard; DNA; 246 BP.
AC
ADK11641;
XX
DT 06-MAY-2004 (first entry)
XX
DE Breast cancer differentially expressed gene product #47.
XX
KW ds; cytostatic; gene therapy; DKFZp5661133 activity inhibitor;
XX
OS breast cancer; differential expression.
XX
OS Homo sapiens.
XX
PN WO2003057925-A1.
XX
PD 17-JUL-2003.
XX
PF 08-JAN-2003; 2003WO-US000657.
XX
PR 08-JAN-2002; 2002US-0345637P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Hansen R;
XX

DR WPI; 2003-577534/54.
XX Inhibiting a cancerous phenotype of a cell, useful for treating breast
PT cancer comprises contacting a cancerous mammalian cell with an agent for
PT inhibition of DRFZp5661l33 activity.
XX Claim 30; SEQ ID NO 47; 257pp; English.
PS The invention relates to a method of inhibiting a cancerous phenotype of
CC a cell comprises contacting a cancerous mammalian cell with an agent for
CC inhibition of DRFZp5661l33 activity. The methods are useful for treating
CC cancer, e.g. breast cancer. This sequence represents a gene product which
CC is differentially expressed in breast cancer cells. The sequence can be
CC used in the method of the invention.
XX
SQ Sequence 246 BP; 77 A; 49 C; 59 G; 61 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 20.8 Length: 246
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-774-176-15 (1-9) x ADK11641 (1-246)
QY 1 LeuilegLYAlaIlePheLeuLeuVal 9
Db 6 CTGATAGCGCTATTTCCTCTGGTT 32

RESULT 2
ADU11677
ID ADU11677 standard; DNA; 475 BP.
XX
AC ADU11677;
XX
DT 27-JAN-2005 (first entry)
XX
DE Solid tumour prognosis gene seqid 2116.
XX
KW cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
OS Unidentified.
XX
FN WO2004097052-A2.
XX
PD 11-NOV-2004.
XX
PF 29-APR-2004; 2004WO-US013587.
XX
PR 29-APR-2003; 2003US-0466067P.
PR 23-JAN-2004; 2004US-0538246P.
XX
PA (AMHP) WYETH.
PA (STRA/) STRAHS A.
XX
PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dorner AJ;
PI
DR WPI; 2004-804779/79.
XX
PT A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
PS Disclosure; Page; 111pp; English.
XX
CC The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially

CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 475 BP; 119 A; 125 C; 117 G; 114 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 43.9 Length: 475
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 13 Gaps: 0

US-10-774-176-15 (1-9) x ADU11677 (1-475)

QY 1 LeuilegLYAlaIlePheLeuLeuVal 9

Db 372 CTGATAGCGCTATTTCCTCTGGTT 398

RESULT 3

AAA27060

ID AAA27060 standard; DNA; 901 BP.

XX

AC AAA27060;

XX

DT 22-AUG-2000 (first entry)

XX

DE Canine 5T4 tumour-associated antigen gene.

XX

KW Canine; TAA; tumour-associated antigen; anti-tumour; cytostatic;
KW immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;
XX ds.

OS Canis sp.

XX

FH Key Location/Qualifiers

FT CDS 1..858

FT /tag= a

FT /product= "5T4 antigen"

FT misc_feature 61..74

FT /tag= b

FT /note= "given in the specification but does not seem to
FT be part of the coding sequence and does not encode any
FT corresponding amino acids"

FT misc_feature 135..146

FT /tag= c

FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 207. .216
 FT /tag= d
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 277. .290
 FT /tag= e
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 351. .361
 FT /tag= f
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 422. .436
 FT /tag= g
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 497. .511
 FT /tag= h
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 572. .583
 FT /tag= i
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 644. .653
 FT /tag= j
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 714. .723
 FT /tag= k
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT 784. .801
 FT /tag= l
 FT /note= "given in the specification but does not seem to
 FT be part of the coding sequence and does not encode any
 FT corresponding amino acids"
 FT
 XX WO200029428-A2.
 XX
 XX 25-MAY-2000.
 XX
 XX 18-NOV-1999; 99WO-GB003859.
 XX
 XX 18-NOV-1998; 98GB-00025303.
 XX 27-JAN-1999; 99GB-00001739.
 XX 30-JUL-1999; 99GB-00017995.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 XX Carroll MW, Myers KA;
 XX
 XX WPI; 2000-387735/33.
 XX P-PSDB; AAY94351.
 XX
 XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte
 XX response useful in vaccinating against and in treating tumors.
 XX
 XX Disclosure; Page 78-79; 79pp; English.
 XX
 XX The present sequence encodes the canine 5T4 tumour-associated antigen
 XX (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in

CC carcinomas but has a highly restricted expression pattern in normal adult
 CC tissues. It appears to be strongly correlated to metastasis in colorectal
 CC and gastric cancer. 5T4 antigen may therefore be useful in tumour
 CC diagnosis, targeting and immunotherapy. Mice in which tumours had been
 CC induced were inoculated with a virus expression vector containing the
 CC human or murine 5T4 gene sequence. The 5T4 antigen was shown to be
 CC effective at eliciting an immunotherapeutic anti-tumour response. Both
 CC the nucleic acid encoding the antigen and the antigen itself can be used
 CC to elicit an immune response, preferably CTL or an antibody response in a
 CC subject
 CC
 XX SQ Sequence 901 BP; 178 A; 246 C; 212 G; 153 T; 0 U; 112 Other;
 Alignment Scores:
 Pred. No.: 91 Length: 901
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 3 Gaps: 0
 US-10-774-176-15 (1-9) x AAA27060 (1-901)
 QY 1 LeuileglyAlaIlePheLeuLeuVal 9
 DB 660 CTGATAGGCGCCATCTTCTACTGTT 686
 RESULT 4
 ABT07721
 ID ABT07721 standard; DNA; 927 BP.
 XX AC ABT07721;
 XX DT 14-NOV-2002 (first entry)
 XX DE Breast cancer-associated gene sequence 29.
 XX Gene; ds; breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.
 XX OS Unidentified.
 XX PN WO200259377-A2.
 XX PD 01-AUG-2002.
 XX PF 24-JAN-2002; 2002WO-US002242.
 XX PR 24-JAN-2001; 2001US-0263965P.
 XX PR 02-FEB-2001; 2001US-0265928P.
 XX PR 09-APR-2001; 2001US-00829472.
 XX PR 09-APR-2001; 2001US-0282698P.
 XX PR 04-MAY-2001; 2001US-0288590P.
 XX PR 29-MAY-2001; 2001US-0294443P.
 XX (BOSB-) EOS BIOTECHNOLOGY INC.
 XX Mack DH, Gish KC, Afar D;
 XX WPI; 2002-583738/62.
 XX N-PSDB; ABJ05564.
 XX
 XX Detecting a breast cancer-associated transcript in a patient's cell,
 XX useful for diagnosing breast cancer, comprises contacting a biological
 XX sample with a polynucleotide that selectively hybridizes with breast
 XX cancer nucleic acids.
 XX
 XX Claim 9; Page 372; 414pp; English.
 XX
 XX The invention comprises a method of detecting a breast cancer-associated
 XX transcript in a cell from a patient. The method of the invention involves
 XX contacting a biological sample from the patient with a nucleotide that
 XX hybridises to one of the 69 breast cancer-associated gene sequences shown

CC in the specification. The method of the invention is useful in the
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are
 CC up or down-regulated in breast cancer cells. Genes identified by the
 CC method of the invention can be used in diagnostic purposes and also as
 CC targets for screening for therapeutic compounds that modulate breast
 CC cancer (e.g. hormones or antibodies). Identification of genes that are
 CC over or under expressed in breast cancer can additionally provide high-
 CC resolution, high-sensitivity datasets which can be used in the areas of
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein
 CC structure and biosensor development. DNA sequences ABT07693 - ABT07761
 CC represent the 69 breast cancer-associated gene sequences of the invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores: 94 Length: 927
 Pred. No.: 40.00 Matches: 9
 Score: 40.00
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-774-176-15 (1-9) x ABT07721 (1-927)

Qy 1 LeulleGlyAlaIlePheLeuLeuVal 9
 |||||
 Db 751 CTGATAGCGCGTATTTCTCTCGTT 777

RESULT 5

ABX76333

ID ABX76333 standard; DNA; 927 BP.

XX AC ABX76333;

XX DT 02-APR-2003 (first entry)

XX DE

XX Lung cancer-associated polynucleotide #197.
 XX

KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX OS

XX WO200286443-A2.

XX PN

XX 31-OCT-2002.

XX PD

XX 18-APR-2002; 2002WO-US012476.

XX PF

XX 18-APR-2001; 2001US-0284770P.

XX PR

XX 10-MAY-2001; 2001US-0290492P.

XX PR

XX 09-NOV-2001; 2001US-0339245P.

XX PR

XX 13-NOV-2001; 2001US-0350666P.

XX PR

XX 29-NOV-2001; 2001US-0334370P.

XX PR

XX 12-APR-2002; 2002US-0372246P.

XX XX

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX PA

XX Aziz N, Murray R;

XX PI

XX WPI; 2003-093161/08.

XX DR

XX P-PSDB; ABUS6604.

XX XX

PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX XX

XX Claim 22; Page 336; 453pp; English.

XX PS

XX XX

XX

CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 94 Length: 927
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-774-176-15 (1-9) x ABX76333 (1-927)

Qy 1 LeulleGlyAlaIlePheLeuLeuVal 9

|||||

Db 751 CTGATAGCGCGTATTTCTCTCGTT 777

RESULT 6

ADB80503

ID ADB80503 standard; DNA; 927 BP.

XX AC ADB80503;

XX DT 04-DEC-2003 (first entry)

XX DE

XX Ovarian cancer-associated transcript #34.

XX KW

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;

XX post-operative chemotherapy; radiation therapy; tumour prognosis;

XX pre-cancerous lesion detection; ds; gene.

XX KW

XX Homo sapiens.

XX OS

XX Key Location/Qualifiers

FT CDS 1..927

FT /*tag= a

XX XX

XX WO2002102235-A2.

XX PN

XX 27-DEC-2002.

XX PD

XX 18-JUN-2002; 2002WO-US019297.

XX PF

XX 18-JUN-2001; 2001US-0299234P.

XX PR

XX 27-AUG-2001; 2001US-0315287P.

XX PR

XX 05-SEP-2001; 2001US-0317544P.

XX PR

XX 13-NOV-2001; 2001US-0350666P.

XX PR

XX 12-APR-2002; 2002US-0372246P.

XX XX

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX PA

XX Mack DH, Gish KC;

XX PI

XX WPI; 2003-167431/16.

XX DR

DR P-PSDB; ADB80504.

XX Detecting an ovarian cancer-associated transcript in a cell from a

PT patient, comprises hybridizing a biological sample from the patient with a

PT polynucleotide that hybridizes to an ovarian cancer gene.

XX

PS Claim 10; Page 297; 332pp; English.

XX

CC The invention relates to a method of detecting an ovarian cancer-

CC associated transcript in a cell from a patient, by contacting a

CC biological sample from the patient with a polynucleotide that selectively

CC hybridizes to a sequence at least 80% identical to any of one of 80

CC nucleic acid sequences given in the specification. The method is useful

CC in diagnosing ovarian cancer and in identifying and using agents and/or

CC targets that inhibit ovarian cancer. The nucleic acid molecule,

CC polypeptide and the antibody may also be used in detecting ovarian

CC cancers, monitoring and early detection of relapse following treatment,

CC monitoring response to therapy, selecting patients for post-operative

CC chemotherapy or radiation therapy, in selecting mode of therapy,

CC determining tumour prognosis, early detection of pre-cancerous lesions,

CC and as vaccines. This sequence corresponds to one of the nucleic acids

CC used for the detection method of the invention.

XX

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	94	Length:	927
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-774-176-15 (1-9) x ADB80503 (1-927)

Qy 1 LeulleGlyAlaIlePheLeuVal 9

Db 751 CTGATAGCGCTATTTCTCTCGTT 777

RESULT 7

ADN38723

ID ADN38723 standard; cDNA; 927 BP.

XX ADN38723;

XX

DT 17-JUN-2004 (first entry)

XX

DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:41.

XX

KW Human; differential expression; cancer; angiogenic disorder;

KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;

KW inflammatory disease; autoimmune disease;

KW retinal neovascularisation syndrome; scarring; uterine fibroid;

KW detection; diagnosis; prognosis; drug screening; drug targeting;

KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;

KW vulnery; gene therapy; vaccine; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003042661-A2.

XX

PD 22-MAY-2003.

XX

XX

PF 13-NOV-2002; 2002WO-US036810.

XX

PR 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.

PR 10-JAN-2002; 2002US-0347349P.

PR 08-FEB-2002; 2002US-035250P.

PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-0368809P.

PR 04-APR-2002; 2002US-0370110P.

PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.

PR 16-JUL-2002; 2002US-0396839P.

PR 22-JUL-2002; 2002US-0397775P.

PR 22-JUL-2002; 2002US-0397845P.

PR 09-SEP-2002; 2002US-0409450P.

XX

PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX

XX WPI; 2003-468649/44.

DR P-PSDB; ADN38724.

XX

PT Determining the presence or absence of a pathological cell in a patient,

PT useful for diagnosing, prognosing or treating cancer, comprises detecting

PT a nucleic acid in a biological sample.

XX

XX Claim 8; SEQ ID NO 41; 1385pp; English.

XX

CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)

CC whose expression is upregulated or downregulated in specific cancers or

CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a

CC patient by detecting a nucleic acid at least 80% identical to those of

CC the invention or by detecting a polypeptide of the invention. The

CC invention also relates to expression vectors and host cells comprising a

CC nucleic acid of the invention; antibodies which specifically bind a

CC polypeptide of the invention; use of such antibodies for drug targeting;

CC and methods of screening for modulators of activity or expression of the

CC polypeptides and nucleic acids. The nucleic acids, polypeptides,

CC antibodies and methods are useful for diagnosing, prognosing and treating

CC cancer and other conditions such as psoriasis, ischaemia, heart disease,

CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal

CC neovascularisation syndromes, scarring and uterine fibroids. They may

CC also be useful in wound healing and in contraception. The present

CC sequence represents a nucleic acid sequence of the invention.

XX

SQ Sequence 927 BP; 187 A; 297 C; 240 G; 203 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	94	Length:	927
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	11	Gaps:	0

US-10-774-176-15 (1-9) x ADN38723 (1-927)

Qy 1 LeulleGlyAlaIlePheLeuVal 9

Db 751 CTGATAGCGCTATTTCTCTCGTT 777

RESULT 8

AAD56198

ID AAD56198 standard; DNA; 973 BP.

XX

XX AAD56198;

XX

DT 07-AUG-2003 (first entry)

XX

DE Human LRRCAPS related DNA #5.

XX

KW Human; p53 pathway; Leucine rich repeat capricious related protein;

KW LRRCAPS; cancer; gene therapy; ds.

XX

OS Homo sapiens.

```

XX PN WO2003035931-A2.
XX XX
XX PD 01-MAY-2003.
XX XX
XX PF 21-OCT-2002; 2002WO-US033540.
XX XX
XX PR 22-OCT-2001; 2001US-0338733P.
XX PR 15-FEB-2002; 2002US-0357600P.
XX PR 01-MAR-2002; 2002US-0361196P.
XX PA (EXEL-) EXELIXIS INC.
XX XX
XX PI Belvin M, Schleithoff L, Plowman GD, Funke RP, Lioubin MN, Li D;
XX PI Francis-Lang H, Friedman L;
XX XX
XX DR WPI; 2003-421410/39.
XX XX
XX PT Identifying a candidate p53 pathway-modulating agent for treating cancer
XX PT comprises contacting an assay system comprising a purified leucine rich
XX PT repeat, capricious related polypeptide or nucleic acid with a test agent.
XX PS Example 5; Page 74-75; 99pp; English.
XX CC The invention relates to a method of identifying a candidate p53 pathway
XX CC modulating agent. The method involves contacting an assay system
XX CC comprising a purified leucine rich repeat, capricious related (LRRCAPS)
XX CC polypeptide or nucleic acid or its fragment with a test agent and
XX CC detecting a test agent-biased activity, where a difference between the
XX CC test agent-biased activity and the reference activity identifies the test
XX CC agent as a candidate p53 pathway modulating agent. The method is useful
XX CC for identifying a candidate p53 pathway-modulating agent for preparing a
XX CC composition for diagnosing or treating cancer. The invention is useful in
XX CC gene therapy. The present sequence is human LRRCAPS related DNA
XX SQ Sequence 973 BP; 203 A; 308 C; 254 G; 208 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 99.3 Length: 973
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-774-176-15 (1-9) x AAD56198 (1-973)
Qy 1 LeulleGlyAlaIlePheLeuVal 9
Db 766 CTGATAGCGCTATTTCTCTCGTT 792

RESULT 9
ABV99349
ID ABV99349 standard; DNA; 1156 BP.
XX AC ABV99349;
XX DT 27-JAN-2003 (first entry)
XX DE Human NOV8a coding sequence.
XX KW Human; anti-HIV; cytostatic; antidiabetic; antiaesthetic; cachexia; AIDS;
KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
KW neurotropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antinfertility; cerebroprotective; gene therapy; NOVA; NOV; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.
XX OS Homo sapiens.
XX XX

```

(CURA-) CURAGEN CORP.

PI Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 PI Voss EA, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 PI Zhong M;
 DR WPI; 2002-732824/79.
 DR P-PSDB; ABP70071.
 XX
 XX New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 PT disorders, and asthma.
 XX
 XX Claim 16; Page 114-115; 619pp; English.
 XX
 XX The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious diseases, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, haematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods
 XX
 SQ Sequence 1156 BP; 228 A; 383 C; 303 G; 242 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 121 Length: 1156
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-15 (1-9) x ABV99349 (1-1156)
 Qy 1 LeulleGlyAlaIlePheLeuLeuVal 9
 Db 982 CTGATAGCGCTATTTCTCTCTGTT 1008
 RESULT 10
 ABK87175
 ID ABK87175 standard; cDNA; 1260 BP.
 XX
 XX ABK87175;
 XX
 XX 07-OCT-2002 (first entry)
 XX
 XX cDNA encoding feline oncofoetal leucine-rich glycoprotein, 5T4.
 XX
 KW Feline; cat; oncofoetal leucine-rich glycoprotein; 5T4; tumour;
 KW cell proliferative disorder; infection; inflammatory condition;
 KW cancer immunotherapy; foetal cell; maternal blood; cytostatic;
 KW foetal abnormality; foetal sex determination; gene; ss.
 XX
 XX Felis sp.
 OS
 XX Key Location/Qualifiers
 FH 1..1260
 FT CDS /tag= a
 FT

FT
 XX /product= "5T4 protein"
 PN WO200238612-A2.
 XX
 PD 16-MAY-2002.
 XX
 XX 13-NOV-2001; 2001WO-GB005004.
 PF
 XX 13-NOV-2000; 2000WO-GB004317.
 PR
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 PA
 XX Myers K, Drury N, Carroll M;
 PI
 XX WPI; 2002-557449/59.
 DR P-PSDB; AAU98694.
 DR
 XX
 XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
 PT polypeptide, useful in preparation of vaccine for treating and/or
 PT preventing cancer in a subject, preferably a dog or cat.
 XX
 XX Claim 4; Page 68; 68pp; English.
 XX
 XX The present invention relates to the isolation of canine and feline
 CC oncofoetal leucine-rich glycoproteins known as 5T4, and the
 CC polynucleotide sequences encoding them. The 5T4 proteins are expressed in
 CC a significant proportion of tumours. The sequences of the invention are
 CC useful in a pharmaceutical composition for the prevention and/or
 CC treatment of tumours or other diseases associated with cell
 CC proliferation, infections, and inflammatory conditions in animals,
 CC preferably dogs or cats. The compositions may also be used for cancer
 CC immunotherapy in these animals. The sequences of the invention may also
 CC be used in diagnostic kits for rapid, reliable, sensitive, and specific
 CC measurement and localisation of 5T4 in extracts of plasma, urine,
 CC tissues, and in cell culture media. Antibodies specific for the 5T4
 CC protein are useful for isolating foetal cells from maternal blood. The
 CC isolation process may form part of a diagnostic method e.g. the foetal
 CC cells may then be subject to biochemical or genetic sampling used for
 CC testing foetal abnormalities, or to determine the sex of the foetus(es).
 CC The present sequence encodes feline 5T4 protein
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 133 Length: 1260
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-774-176-15 (1-9) x ABK87175 (1-1260)
 Qy 1 LeulleGlyAlaIlePheLeuLeuVal 9
 Db 1090 CTGATAGGTGCCATTTTCTTACTGTT 1116
 RESULT 11
 ADB97513
 ID ADB97513 standard; DNA; 1260 BP.
 XX
 XX ADB97513;
 XX
 XX 04-DEC-2003 (first entry)
 DT
 XX Feline 5T4 antigen DNA.
 DE
 XX Major Histocompatibility Complex class I peptide epitope; MHC;
 KW 5T4 antigen; 5T4 epitope; polypeptide string; vaccine; T cell;
 KW cytostatic; cancer; feline; gene; ds.
 XX
 XX Unidentified.
 OS
 XX

PH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"
 XX
 XX WO2003068816-A1.
 XX
 PD 21-AUG-2003.
 XX
 PF 13-FEB-2003; 2003WO-GB000670.
 XX
 PR 13-FEB-2002; 2002GB-00003419.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 XX Carroll M, Kingman S, Redchenko I;
 XX
 DR WPI; 2003-637141/60.
 DR P-PSDB; ADB97520.
 XX
 XX New major histocompatibility complex class I peptide epitopes from human
 PT 5T4 tumor-associated antigen, useful for preventing and/or treating a
 PT disease, particularly cancer.
 XX
 XX Disclosure; Page 67; 73pp; English.
 XX
 CC The invention relates to a novel Major Histocompatibility Complex (MHC)
 CC class I peptide epitope of the 5T4 antigen. The invention further
 CC provides a polypeptide string comprising the 5T4 epitope; a nucleic acid
 CC sequence encoding the 5T4 epitope or a polypeptide string of the 5T4
 CC epitope; a vector system capable of delivering the 5T4 epitope nucleic
 CC acid to a cell; a cell pulsed with the 5T4 epitope, a polypeptide of the
 CC 5T4 epitope, its encoding nucleic acid, or the vector system; a vaccine
 CC comprising the above; a method for treating and/or preventing a disease
 CC in a subject by administering the vaccine; an agent capable of binding
 CC specifically to the 5T4 epitope and/its encoding nucleic acid; a method
 CC comprising detecting the presence of the 5T4 epitope or its encoding
 CC nucleic acid in a subject; and a T cell line or clone capable of
 CC specifically recognising the 5T4 epitope in conjunction with an MHC class
 CC I molecule. The 5T4 epitope has cytostatic activity. The vaccine
 CC comprising the 5T4 epitope or its encoding nucleic acid and the vector
 CC system or cell is useful in the prevention and/or treatment of a disease,
 CC particularly cancer. The detection method is useful for diagnosing or
 CC monitoring the progression of a cancerous disease, and for detecting the
 CC presence of the 5T4 epitope or its nucleic acid. The T cell line or clone
 CC is useful in the manufacture of a medicament for treating and/or
 CC preventing a disease. This polynucleotide sequence represents the feline
 CC 5T4 antigen coding DNA of the invention.
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 XX
 Alignment Scores:
 Pred. No.: 133 Length: 1260
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-15 (1-9) x ADB97513 (1-1260)
 QY 1 LeuileGlyAlaIlePheLeuVal 9
 Db 1090 CTGATAGTGCCATTTTCTTACTGGTT 1116
 RESULT 12
 ADB97452
 ID ADB97452 standard; DNA; 1260 BP.
 XX
 AC ADB97452;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 KW Human 5T4 tumour-associated antigen; anti-tumour; cytostatic;
 immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;

DE DNA encoding feline 5T4 protein.
 XX
 XX Gene; ds; feline; Major Histocompatibility Complex class II; MHC;
 KW epitope; 5T4 antigen; vaccine; T-cell; cytostatic; cancer; gene therapy.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1260
 FT /*tag= a
 FT /product= "Feline 5T4 antigen protein"
 XX
 PN WO2003068815-A2.
 XX
 PD 21-AUG-2003.
 XX
 PF 13-FEB-2003; 2003WO-GB000618;
 XX
 PR 13-FEB-2002; 2002GB-00003420.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 PA
 XX Carroll M, Harrop R, Kingsman S;
 XX
 DR WPI; 2003-663795/62.
 DR P-PSDB; ADB97455.
 XX
 XX New Major Histocompatibility Complex class II peptide epitope of 5T4,
 PT useful for manufacturing a medicament for diagnosing, preventing and/or
 PT treating a disease, e.g. cancer.
 XX
 PS Disclosure; Page 49; 63pp; English.
 XX
 CC The invention relates to a Major Histocompatibility Complex (MHC) class
 CC II peptide epitope of the 5T4 antigen. The vaccine or T-cell line or
 CC clone has a cytostatic activity, as it is useful in manufacturing a
 CC medicament for preventing and/or treating a disease, particularly cancer.
 CC The methods are useful for detecting T-cells capable of specifically
 CC recognising a peptide epitope in conjunction with an MHC molecule, for
 CC diagnosing or monitoring the progression of a cancerous disease, or for
 CC detecting the presence of a peptide or nucleic acid using an agent. The
 CC MHC class II peptide epitope of the invention can be used in gene therapy
 CC or as part of a vaccine. This polynucleotide sequence represents the DNA
 CC coding for the feline 5T4 protein.
 XX
 SQ Sequence 1260 BP; 221 A; 451 C; 346 G; 241 T; 0 U; 1 Other;
 XX
 Alignment Scores:
 Pred. No.: 133 Length: 1260
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-774-176-15 (1-9) x ADB97452 (1-1260)
 QY 1 LeuileGlyAlaIlePheLeuVal 9
 Db 1090 CTGATAGTGCCATTTTCTTACTGGTT 1116
 RESULT 13
 AAA27058
 ID AAA27058 standard; DNA; 1263 BP.
 XX
 AC AAA27058;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human 5T4 tumour-associated antigen gene.
 KW Human; TAA; tumour-associated antigen; anti-tumour; cytostatic;
 immunostimulant; vaccine; carcinoma; colorectal cancer; gastric cancer;

KW ds.
XX Homo sapiens.
XX WO200029428-A2.
XX 25-MAY-2000.
XX 18-NOV-1999; 99WO-GB003859.
XX 18-NOV-1998; 98GB-00025303.
XX 27-JAN-1999; 99GB-00001739.
XX 30-JUL-1999; 99GB-00017995.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX Carroll MW, Myers KA;
XX WPI; 2000-387735/33.
XX Tumor associated antigen, 5T4 capable of eliciting cytotoxic T-lymphocyte response useful in vaccinating against and in treating tumors.
XX Example 2; Page 78; 79pp; English.
XX The present sequence encodes the human 5T4 tumour-associated antigen (TAA). The TAA 5T4 is a glycoprotein which is widely expressed in carcinomas but has a highly restricted expression pattern in normal adult tissues. It appears to be strongly correlated to metastasis in colorectal and gastric cancer. 5T4 antigen may therefore be useful in tumour diagnosis, targeting and immunotherapy. Mice in which tumours had been induced were inoculated with a virus expression vector containing the present sequence. The 5T4 antigen was shown to be effective at eliciting an immunotherapeutic anti-tumour response. Both the nucleic acid encoding the antigen and the antigen itself can be used to elicit an immune response, preferably CTL or an antibody response in a subject
XX
SQ Sequence 1263 BP; 230 A; 428 C; 349 G; 256 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 134 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
US-10-774-176-15 (1-9) x AAA27058 (1-1263)
Qy 1 LeulleGlyAlaIlePheLeuVal 9
Db 1093 CTGATAGCGCTATTTCTCTCTGTT 1119
RESULT 14
AAF89736
ID AAF89736 standard; DNA; 1263 BP.
XX AAF89736;
XX 23-JUL-2001 (first entry)
XX Nucleotide sequence of canine 5T4 protein.
XX Single chain antibody; ScFv; inflammatory disease; arthritis; cancer; hypersensitivity; autoimmune disease; central nervous system disorder; Parkinson's disease; periodontal disease; cardiovascular disease; Helicobacter-related disease; immune disorder; ss.
XX Canis sp.
XX Key Location/Qualifiers
XX CDS 1..1263

FT /*tag= a
FT /product= "5T4"
XX WO200136486-A2.
XX 25-MAY-2001.
XX 13-NOV-2000; 2000WO-GB004317.
XX 18-NOV-1999; 99WO-GB003859.
XX 15-FEB-2000; 2000GB-00003527.
XX 02-MAR-2000; 2000GB-00005071.
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard FM;
XX Myers KA;
XX WPI; 2001-343805/36.
XX P-PSDB; AAB83839.
XX Use of single chain antibody capable of recognizing a disease associated molecule for manufacturing a medicament for preventing and/or treating a disease condition associated with disease associated molecule.
XX Disclosure; Fig 26; 118pp; English.
XX The specification describes the use of a single chain antibody (ScFv), which is capable of recognizing a disease associated molecule in the manufacture of a medicament for the prevention and treatment of a disease condition. The ScFv antibody is useful in the manufacture of a medicament, for affecting a disease in vivo, for preparing a pharmaceutical composition, for in vivo imaging and/or for adjuvant treatment of a disease. The ScFv antibody is also useful for treating inflammatory diseases including arthritis, hypersensitivity, autoimmune diseases, cancers, central nervous system disorders including Parkinson's disease, periodontal diseases, cardiopulmonary diseases, cardiovascular diseases, gastrointestinal disorders, infections, diabetes, Helicobacter-related diseases, and other immune disorders. The present sequence encodes a 5T4 protein, which is used to produce ScFv of the invention
XX
SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 134 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 4 Gaps: 0
US-10-774-176-15 (1-9) x AAF89736 (1-1263)
Qy 1 LeulleGlyAlaIlePheLeuVal 9
Db 1093 CTGATAGCGCCATCTTCTACTGTT 1119
RESULT 15
ABK87174
ID ABK87174 standard; cDNA; 1263 BP.
XX ABK87174;
XX 07-OCT-2002 (first entry)
XX cDNA encoding canine oncofoetal leucine-rich glycoprotein, 5T4.
XX Canine; dog; oncofoetal leucine-rich glycoprotein; 5T4; tumour; cell proliferative disorder; infection; inflammatory condition; cancer immunotherapy; foetal cell; maternal blood; cytostatic; foetal abnormality; foetal sex determination; gene; ss.
XX Canis sp.

XX Key Location/Qualifiers
PH 1..1263
FT CDS
FT /*tag= a
FT /product= "5T4 protein"
XX
XX WO200238612-A2.
XX
XX 16-MAY-2002.
XX
XX 13-NOV-2001; 2001WO-GB005004.
XX
XX 13-NOV-2000; 2000WO-GB004317.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Myers K, Drury N, Carroll M;
XX
XX WPI; 2002-557449/59.
XX P-PSDB; AAU98693.
XX
XX Novel canine or feline 5T4 polypeptide and polynucleotides encoding the
XX polypeptide, useful in preparation of vaccine for treating and/or
XX preventing cancer in a subject, preferably a dog or cat.
XX
XX Claim 1; Page 67; 68pp; English.
XX
XX The present invention relates to the isolation of canine and feline
XX oncofoetal leucine-rich glycoproteins known as 5T4, and the
XX polynucleotide sequences encoding them. The 5T4 proteins are expressed in
XX a significant proportion of tumours. The sequences of the invention are
XX useful in a pharmaceutical composition for the prevention and/or
XX treatment of tumours or other diseases associated with cell
XX proliferation, infections, and inflammatory conditions in animals,
XX preferably dogs or cats. The compositions may also be used for cancer
XX immunotherapy in these animals. The sequences of the invention may also
XX be used in diagnostic kits for rapid, reliable, sensitive, and specific
XX measurement and localisation of 5T4 in extracts of plasma, urine,
XX tissues, and in cell culture media. Antibodies specific for the 5T4
XX protein are useful for isolating foetal cells from maternal blood. The
XX isolation process may form part of a diagnostic method e.g. the foetal
XX cells may then be subject to biochemical or genetic sampling used for
XX testing foetal abnormalities, or to determine the sex of the foetus(es).
XX The present sequence encodes canine 5T4 protein
XX
SQ Sequence 1263 BP; 200 A; 469 C; 377 G; 217 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 134 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-774-176-15 (1-9) x ABK87174 (1-1263)
QY 1 LeuIleGlyAlaIlePheLeuVal 9
Db 1093 CTGATAGGCCCATCTTCTACTGTT 1119

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: May 27, 2006, 09:51:03 ; Search time 3358.6 Seconds
(without alignments)
257.039 Million cell updates/sec

Title: US-10-774-176-15
Perfect score: 40
Sequence: 1 LIGAIFLLV 9

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Ddelop 6.0 , Ddelxt 7.0

Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB.spool/US10774176/runat.26052006.091443.24987/app.query.fasta_1
-DB=GenEmbl -OPMT=faetap -SUPFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes04
-USER=US10774176 @CGN 1 1 7524 @runat.26052006.091443.24987 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WEN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sta.*
8: gb_sy.*
9: gb_un.*
10: gb_vl.*
11: gb_ov.*
12: gb_htg.*
13: gb_in.*
14: gb_on.*
15: gb_pa.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	290	2	CQ687716 Sequence
2	40	100.0	475	2	CQ920916 Sequence
3	40	100.0	901	2	BD249733 Polypepti

4	40	100.0	901	2	AX025013	AX025013 Sequence
5	40	100.0	901	2	AX316088	AX316088 Sequence
6	40	100.0	927	2	AX829164	AX829164 Sequence
7	40	100.0	1156	2	DD161112	DD161112 Novel Ant
8	40	100.0	1260	2	AX467373	AX467373 Sequence
9	40	100.0	1260	2	AX821533	AX821533 Sequence
10	40	100.0	1260	2	AX821548	AX821548 Sequence
11	40	100.0	1263	2	BD249731	BD249731 Polypepti
12	40	100.0	1263	2	AX025011	AX025011 Sequence
13	40	100.0	1263	2	AX149553	AX149553 Sequence
14	40	100.0	1263	2	AX316086	AX316086 Sequence
15	40	100.0	1263	2	AX467371	AX467371 Sequence
16	40	100.0	1281	2	BD249732	BD249732 Polypepti
17	40	100.0	1281	2	AX025012	AX025012 Sequence
18	40	100.0	1281	2	AX316087	AX316087 Sequence
19	40	100.0	2053	2	CQ731678	CQ731678 Sequence
20	40	100.0	2053	2	DD174290	DD174290 METHODS O
21	40	100.0	2053	5	HS5740A	229083 Homo sapien
22	40	100.0	2333	6	AF063939	AF063939 Rattus no
23	40	100.0	2359	2	BD127282	BD127282 Primer fo
24	40	100.0	2359	2	CQ782724	CQ782724 Sequence
25	40	100.0	2359	5	AK074786	AK074786 Homo sapi
26	40	100.0	2361	2	AX961916	AX961916 Sequence
27	40	100.0	2361	2	BD127283	BD127283 Primer fo
28	40	100.0	2361	2	CQ782726	CQ782726 Sequence
29	40	100.0	2361	5	AK074790	AK074790 Homo sapi
30	40	100.0	2361	6	BC087011	BC087011 Rattus no
31	40	100.0	2379	5	BC037161	BC037161 Homo sapi
32	40	100.0	2423	6	BC058198	BC058198 Mus muscu
33	40	100.0	2557	2	AX961912	AX961912 Sequence
34	40	100.0	2557	2	AX961914	AX961914 Sequence
35	40	100.0	2714	5	AB168308	AB168308 Macaca fa
36	40	100.0	5551	6	HSA012159	AJ012159 Homo sapi
37	40	100.0	7942	6	MMU012160	AJ012160 Mus muscu
38	40	100.0	121909	5	HSJ492P14	AL121977 Human DNA
39	40	100.0	167046	6	AC158516	AC158516 Mus muscu
40	40	100.0	210237	12	AC128294	AC128294 Rattus no
41	40	100.0	239076	12	AC106962	AC106962 Rattus no
42	40	100.0	243801	12	AC172361	AC172361 Bos tauru
43	40	100.0	302050	15	AL935256	AL935256 Lactobaci
44	39	97.5	137667	14	AC145041	AC145041 Macropus
45	39	97.5	142071	12	AC149958	AC149958 Strongylo

ALIGNMENTS

RESULT 1	CQ687716	Sequence 32642 from Patent WO02070737.	290 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ687716	Sequence 32642 from Patent WO02070737.				
DEFINITION	CQ687716	Sequence 32642 from Patent WO02070737.				
ACCESSION	CQ687716	Sequence 32642 from Patent WO02070737.				
VERSION	CQ687716.1	GI:42218962				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1	Liew,C.C., Marshall,W.E. and Zhang,H.				
AUTHORS		Compositions and methods relating to osteoarthritis				
TITLE		Patent: WO 02070737-A 32642 12-SEP-2002;				
JOURNAL		Chondrogene Inc. (CA)				
FEATURES	source	Location/Qualifiers				
		1..290				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				

ORIGIN

Alignment Scores:	1.09	Length:	290
Pred. No.:	40.00	Matches:	9

DB: 2 Gaps: 0

US-10-774-176-15 (1-9) x AX025013 (1-901)

Qy 1 Leu1leGlyAla1lePheLeuVal 9
 Db 660 CTGATAGGCGCATCTTCTACTGTT 686

RESULT 5
 AX316088
 LOCUS AX316088 901 bp DNA linear PAT 14-DEC-2001
 DEFINITION Sequence 3 from Patent EP1160323.
 ACCESSION AX316088
 VERSION AX316088.1 GI:178993280
 KEYWORDS
 SOURCE Canis sp.
 ORGANISM Canis sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 Canis.

REFERENCE 1
 AUTHORS Carroll, M.W. and Myers, K.A.
 TITLE 5t4 tumour-associated antigen for use in tumour immunotherapy
 JOURNAL Patent: EP 1160323-A 3 05-DEC-2001;
 Oxford Biomedica (UK) Limited (GB)

FEATURES
 source
 1..901
 /organism="Canis sp."
 /mol_type="unassigned DNA"
 /db_xref="taxon:9616"

ORIGIN

Alignment Scores:
 Pred. No.: 3.54 Length: 901
 Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

US-10-774-176-15 (1-9) x AX316088 (1-901)

Qy 1 Leu1leGlyAla1lePheLeuVal 9
 Db 660 CTGATAGGCGCATCTTCTACTGTT 686

RESULT 6
 AX829164
 LOCUS AX829164 927 bp DNA linear PAT 12-DEC-2003
 DEFINITION Sequence 57 from Patent WO02059377.
 ACCESSION AX829164
 VERSION AX829164.1 GI:39838931
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE 1
 AUTHORS Mack, D.H., Gish, K.C. and Afar, D.
 TITLE Methods of diagnosis of breast cancer, compositions and methods of
 screening for modulators of breast cancer
 JOURNAL Patent: WO 02059377-A 57 01-AUG-2002;
 EOS Biotechnology, Inc. (US)

FEATURES
 source
 1..927
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 3.65 Length: 927

Score: 40.00 Matches: 9
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

US-10-774-176-15 (1-9) x AX829164 (1-927)

Qy 1 Leu1leGlyAla1lePheLeuVal 9
 Db 751 CTGATAGGCGCATTTTCTCTCTGTT 777

RESULT 7
 DD161112
 LOCUS DD161112 1156 bp DNA linear PAT 23-NOV-2005
 DEFINITION Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
 Encoding The Antigens, and Methods of Use.
 ACCESSION DD161112
 VERSION DD161112.1 GI:83967439
 KEYWORDS JP 2005508604-A/23.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 1156)
 AUTHORS Padigar, M., Shenoy, S.G., Pochart, P.F., Kekuda, R., Gusev, V.V.,
 Zhong, M., Jr, R.J.T., Casman, S.J., Li, L., Miller, C.E.,
 Patturajan, M., Anderson, D.W., Malyankar, U.M., Voss, E.Z.,
 Spaderna, S.K., Gorman, L., Spytek, K.A., Liu, X., Burgess, C.E.,
 Pena, C.E.A., Gerlach, V., Smithson, G., Mezes, P.D., Rastelli, L.,
 Boldog, F.L., Guo, X., Vernet, C.A.M., Gangolli, E.A., Tchernev, V.T.
 and Zerhusen, B.D.
 TITLE Novel Antibodies that Bind to Antigenic Polypeptides, Nucleic Acids
 Encoding The Antigens, and Methods of Use
 JOURNAL Patent: JP 2005508604-A 23 07-APR-2005;
 Muralidhara Padigar, Suresh Shenoy, Remesh Kekuda, Vladimir Gusev,
 Pascale Pochart, Mei Zhong, Luca Rastelli, Peter Mezes, Glennda
 Smithson, Xiaojia Guo, Valerie Gerlach, Stacie Casman, Ferenc
 Boldog, Li Li, Bryan Zerhusen, Velizar Tchernev, Esha Gangolli, Corine
 Vernet, Carol Pena, Catherine Burgess, Xiaohong Liu, Kimberly
 Spytek, Linda Gorman, Steven Spaderna, Edward Voss, Uriel
 Malyankar, David Anderson, Meera Patturajan, Charles Miller, Raymond J
 Taupier Jr

COMMENT
 OS Homo sapiens
 PN JP 2005508604-A/23
 PD 07-APR-2005
 PP 08-MAR-2002 JP 2002571827
 PR 19-JUN-2001 US 60/299310, 18-JUN-2001 US 60/299027, PR
 31-MAY-2001 US 60/294889, 31-MAY-2001 US 60/294899, PR
 30-MAY-2001 US 60/294485, 09-MAR-2001 US 60/274849, PR
 13-MAR-2001 US 60/275579, 03-MAY-2001 US 60/288342, PR
 02-MAR-2001 US 60/288066, 30-APR-2001 US 60/287424, PR
 13-APR-2001 US 60/283675, 04-APR-2001 US 60/281194, PR
 02-APR-2001 US 60/280802, 02-APR-2001 US 60/280822, PR
 03-MAY-2001 US 60/288528, 15-MAY-2001 US 60/291190, PR
 12-MAY-2001 US 60/291240, 16-MAY-2001 US 60/291099, PR
 16-MAR-2001 US 60/275235, 08-MAR-2001 US 60/274101, PR
 08-MAR-2001 US 60/274281, 08-MAR-2001 US 60/274322, PR
 08-MAR-2001 US 60/274194, 02-APR-2001 US 60/280900, PR
 30-MAR-2001 US 60/280233, 30-MAR-2001 US 60/279995, PR
 20-MAR-2001 US 60/277327, 20-MAR-2001 US 60/277338, PR
 19-MAR-2001 US 60/276994, 16-MAR-2001 US 60/276776, PR
 14-MAR-2001 US 60/276000, 13-MAR-2001 US 60/275601, PR
 13-MAR-2001 US 60/275578, 20-MAR-2001 US 60/277239, PR
 20-MAR-2001 US 60/277321, 21-MAR-2001 US 60/277791, PR
 22-MAR-2001 US 60/277833, 23-MAR-2001 US 60/278152, PR
 26-MAR-2001 US 60/278894, 27-MAR-2001 US 60/278999, PR
 27-MAR-2001 US 60/279036, 28-MAR-2001 US 60/279344, PR
 19-JUN-2001 US 60/299303, 10-JUL-2001 US 60/304354, PR
 31-JUL-2001 US 60/309198, 03-DEC-2001 US 60/337426, PR
 03-DEC-2001 US 60/338092, 21-NOV-2001 US 60/332094, PR
 14-NOV-2001 US 60/333272, 14-NOV-2001 US 60/332271, PR

```

14-NOV-2001 US 60/332272,14-NOV-2001 US 60/332172, PR
14-NOV-2001 US 60/333184,31-OCT-2001 US 60/335301, PR
18-OCT-2001 US 60/330380,27-SEP-2001 US 60/325430, PR
27-SEP-2001 US 60/325681,12-SEP-2001 US 60/318770, PR
10-SEP-2001 US 60/318462,03-JAN-2002 US 60/345705, PR
04-DEC-2001 US 60/337185,08-MAR-2002 US 10/093463, PR
16-AUG-2001 US 60/312903
PI muralidhara padigaru,suresh g shenoy,pascale f-g pochart, PI
remesh kekuda,
PI vladimir y gusev,mei zhong,raymond j taupier jr,etacie j PI
casman,li li,
PI charles e miller,meera patturajan,david w anderson,uriel m PI
PI edward z voas,steven k spaderna,linda gorman,kimberly PI a
PI xiaohong liu,catherine e burgess,carol e a pena,valerie PI
spytek,
PI glennnda smithson,peter d mezes,luca rastelli,ferenc l boldog,
gerlach,
PI xiaojia guo,
PI corine a m vernet,esha a gangolli,velizar t tchernev,bryan d
PI zerhusen
CC
FH Key Location/Qualifiers
FT CDS (1)..(1150).
FEATURES
    source
        1..1156
        Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.6 Length: 1156
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-15 (1-9) x DD161112 (1-1156)
QY 1 LeuIlleGlyAlaIlePheLeuLeuVal 9
Db 982 CTGATAGCGCGTATTTCTCTCTGGTT 1008
RESULT 8
AX467373
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE Felis sp.
ORGANISM Felis sp.
REFERENCE
    1 Myers,K., Drury,N. and Carroll,M.
    Polyptide
    Patent: WO 0238612-A 3 16-MAY-2002;
    Oxford Biomedica (UK) Limited (GB)
FEATURES
    source
        1..1260
        Location/Qualifiers
            /organism="Felis sp."
            /mol_type="unassigned DNA"
            /db_xref="taxon:9687"
ORIGIN
Alignment Scores:
Pred. No.: 5.03 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
14-NOV-2001 US 60/332272,14-NOV-2001 US 60/332172, PR
14-NOV-2001 US 60/333184,31-OCT-2001 US 60/335301, PR
18-OCT-2001 US 60/330380,27-SEP-2001 US 60/325430, PR
27-SEP-2001 US 60/325681,12-SEP-2001 US 60/318770, PR
10-SEP-2001 US 60/318462,03-JAN-2002 US 60/345705, PR
04-DEC-2001 US 60/337185,08-MAR-2002 US 10/093463, PR
16-AUG-2001 US 60/312903
PI muralidhara padigaru,suresh g shenoy,pascale f-g pochart, PI
remesh kekuda,
PI vladimir y gusev,mei zhong,raymond j taupier jr,etacie j PI
casman,li li,
PI charles e miller,meera patturajan,david w anderson,uriel m PI
PI edward z voas,steven k spaderna,linda gorman,kimberly PI a
PI xiaohong liu,catherine e burgess,carol e a pena,valerie PI
spytek,
PI glennnda smithson,peter d mezes,luca rastelli,ferenc l boldog,
gerlach,
PI xiaojia guo,
PI corine a m vernet,esha a gangolli,velizar t tchernev,bryan d
PI zerhusen
CC
FH Key Location/Qualifiers
FT CDS (1)..(1150).
FEATURES
    source
        1..1156
        Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.6 Length: 1156
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-15 (1-9) x DD161112 (1-1156)
QY 1 LeuIlleGlyAlaIlePheLeuLeuVal 9
Db 982 CTGATAGCGCGTATTTCTCTCTGGTT 1008
RESULT 8
AX467373
LOCUS AX467373 1260 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0238612.
ACCESSION AX467373
VERSION AX467373.1 GI:21900603
KEYWORDS
SOURCE Felis sp.
ORGANISM Felis sp.
REFERENCE
    1 Myers,K., Drury,N. and Carroll,M.
    Polyptide
    Patent: WO 0238612-A 3 16-MAY-2002;
    Oxford Biomedica (UK) Limited (GB)
FEATURES
    source
        1..1260
        Location/Qualifiers
            /organism="Felis sp."
            /mol_type="unassigned DNA"
            /db_xref="taxon:9687"
ORIGIN
Alignment Scores:
Pred. No.: 5.03 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

```

```

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-15 (1-9) x AX467373 (1-1260)
QY 1 LeuIlleGlyAlaIlePheLeuLeuVal 9
Db 1090 CTGATAGTGCCATTTTCTTACTGGTT 1116
RESULT 9
AX821533
LOCUS AX821533 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068816.
ACCESSION AX821533
VERSION AX821533.1 GI:39724929
KEYWORDS
SOURCE Felis catus (cat)
ORGANISM Felis catus
REFERENCE
    1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
    Felinae; Felis.
    Carroll,M.M., Kingsman,S.M. and Redchenko,I.M.
    MHC class I peptide epitopes from the human St4 tumor-associated
    antigen
    Patent: WO 03068816-A 1 21-AUG-2003;
    Oxford Biomedica (UK) Limited (GB)
FEATURES
    source
        1..1260
        Location/Qualifiers
            /organism="Felis catus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 5.03 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-774-176-15 (1-9) x AX821533 (1-1260)
QY 1 LeuIlleGlyAlaIlePheLeuLeuVal 9
Db 1090 CTGATAGTGCCATTTTCTTACTGGTT 1116
RESULT 10
AX821548
LOCUS AX821548 1260 bp DNA linear PAT 10-DEC-2003
DEFINITION Sequence 1 from Patent WO03068815.
ACCESSION AX821548
VERSION AX821548.1 GI:39724930
KEYWORDS
SOURCE Felis catus (cat)
ORGANISM Felis catus
REFERENCE
    1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
    Felinae; Felis.
    Carroll,M.O., Harrop,R.O. and Kingsman,S.O.
    MHC class II peptide epitope of St4 antigen
    Patent: WO 03068815-A 1 21-AUG-2003;
    Oxford Biomedica (UK) Limited (GB)
FEATURES
    source
        1..1260
        Location/Qualifiers
            /organism="Felis catus"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9685"
ORIGIN
Alignment Scores:
Pred. No.: 5.03 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

```

Alignment Scores:
Pred. No.: 5.03 Length: 1260
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-15 (1-9) x AX821548 (1-1260)

Qy 1 LeuileGlyAlaIlePheLeuLeuVal 9
|||||
Db 1090 CTGATAGGCGCATTTTCTTACTGGTT 1116

RESULT 11
LOCUS BD249731 1263 bp DNA linear PAT 17-JUL-2003
DEFINITION Polypeptide.
ACCESSION BD249731
VERSION BD249731.1 GI:33059501
KEYWORDS JP 2002530060-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 1263)
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL OXFORD BIOMEDICA LTD
COMMENT OS Homo sapiens (human)
PN JP 2002530060-A/1
PD 17-SEP-2002
PR 18-NOV-1999 JP 2000582415
PR 18-NOV-1998 GB 9825303.2,27-JAN-1999 GB 9901739.4 PR
30-JUL-1999 GB 9917995.4
PI MILES WILLIAM CARROLL,KEVIN ALAN MYERS
PC C12N15/09,A61K39/00,A61K48/00,A61P35/00,C07K7/06,C07K14/065,
PC C07K19/00,
PC C12N15/00
CC Polypeptide
FH Key Location/Qualifiers
FT source 1..1263
FT /organism="Homo sapiens (human)".

FEATURES
source Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5.04 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-15 (1-9) x BD249731 (1-1263)

Qy 1 LeuileGlyAlaIlePheLeuLeuVal 9
|||||
Db 1093 CTGATAGGCGCATTTTCTTCTCTGGTT 1119

RESULT 12
LOCUS AX025011 1263 bp DNA linear PAT 15-SEP-2000
DEFINITION Sequence 1 from Patent WO029428.
ACCESSION AX025011
VERSION AX025011.1 GI:10184932
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Carroll,M.W. and Myers,K.A.
TITLE Polypeptide
JOURNAL Patent: WO 0029428-A 1 25-MAY-2000;
CARROLL MILES WILLIAM (GB) ; MYERS KEVIN ALAN (GB) ; OXFORD
BIOMEDICA LTD (GB)
FEATURES
source Location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5.04 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-15 (1-9) x AX025011 (1-1263)

Qy 1 LeuileGlyAlaIlePheLeuLeuVal 9
|||||
Db 1093 CTGATAGGCGCATTTTCTTCTCTGGTT 1119

RESULT 13
LOCUS AX149553 1263 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 14 from Patent WO0136486.
ACCESSION AX149553
VERSION AX149553.1 GI:14347991
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Kingsman,A.O., Kingsman,S.M., Bebbington,C.R., Carroll,M.W.,
Ellard,F.M. and Myers,K.A.
TITLE Antibodies
JOURNAL Patent: WO 0136486-A 14 25-MAY-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source Location/Qualifiers
1..1263
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="514"

ORIGIN
Alignment Scores:
Pred. No.: 5.04 Length: 1263
Score: 40.00 Matches: 9
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-774-176-15 (1-9) x AX149553 (1-1263)

Qy 1 LeuileGlyAlaIlePheLeuLeuVal 9
|||||
Db 1093 CTGATAGGCGCATTTTCTTCTCTGGTT 1119

RESULT 14
LOCUS AX316086 1263 bp DNA linear PAT 14-DEC-2001

Search completed: May 27, 2006, 19:35:21
Job time : 3359.6 secs

DEFINITION Sequence 1 from Patent EP1160323.
ACCESSION AX316086
VERSION AX316086.1 GI:17899278
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Carroll, M.W. and Myers, K.A.
TITLE S4 tumour-associated antigen for use in tumour immunotherapy
JOURNAL Patent: EP 1160323-A 1 05-DEC-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1263
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:			
Pred. No.:	5.04	Length:	1263
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-774-176-15 (1-9) x AX316086 (1-1263)

Qy 1 LeuileGlyAlaIlePheLeuVal 9

Db 1093 CTGATAGCGCTATTTCCTCTCTGTT 1119

RESULT 15

AX467371
LOCUS AX467371 1263 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0238612.
ACCESSION AX467371
VERSION AX467371.1 GI:21900602
KEYWORDS
SOURCE Canis sp.
ORGANISM Canis sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.

REFERENCE 1

AUTHORS Myers, K., Drury, N. and Carroll, M.
TITLE Polypeptide
JOURNAL Patent: WO 0238612-A 1 16-MAY-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source 1..1263
/organism="Canis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:9616"

ORIGIN

Alignment Scores:			
Pred. No.:	5.04	Length:	1263
Score:	40.00	Matches:	9
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-10-774-176-15 (1-9) x AX467371 (1-1263)

Qy 1 LeuileGlyAlaIlePheLeuVal 9

Db 1093 CTGATAGCGCCATCTCTCTCTGTT 1119